

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE BIOLOGIA VEGETAL



## **Caracterização de Amostras de Mel por *Next Generation Sequencing***

**Mestrado em Biologia Molecular e Genética**

Inês Maria Pinto Mateus Valbom

Dissertação orientada por:  
Sandra Mourinha Chaves  
Maria Helena Machado Trindade

2016







*"The process of scientific discovery is, in effect, a  
continual flight from wonder."*

*Albert Einstein*



## Agradecimentos

Desde já, gostaria de expressar a minha gratidão a todos que contribuíram para este trabalho e que me acompanharam durante este importante percurso.

Para começar, gostaria de agradecer à minha orientadora Professora Sandra Chaves por me acolher na sua equipa, pelo apoio, enorme disponibilidade e por todo o conhecimento que partilhou comigo ao longo deste ano. Muito obrigada pela oportunidade.

A todos os meus colegas da Biopremier por todos os momentos que passámos juntos e pelos laços de amizade criados.

À Professora Teresa Crespo e Frédéric Gaspar por acompanharem este projeto de perto e por todas as sugestões dadas.

À Professora Maria Helena Trindade os meus agradecimentos pela leitura atenta e pelas sugestões apresentadas.

Ao Professor Rogério Tenreiro por partilhar comigo o seu conhecimento sobre análise multivariada.

Aos meus amigos e colegas de casa, Maria, Pedro e Vítor, por todos os momentos que partilhámos ao longo deste ano. Vocês fazem-me sentir em “casa”.

À Carolina, pela verdadeira e longa amizade.

Ao Tiago, por me mostrar que mesmo nos momentos mais sombrios, existe sempre um céu bem azul.

E porque os últimos são sempre os primeiros:

À minha mãe, obrigada por tudo. Obrigada por todas as oportunidades proporcionadas. Obrigada por me guiares diariamente e por todos os valores que me transmitiste ao longo dos anos. Obrigada por me orientares e amparares nos momentos mais difíceis, não só durante o meu percurso académico, mas sim neste caminho que é a vida. És o meu maior exemplo e orgulho. Obrigada por todo o teu amor. O meu nunca te vai faltar. Mãe, obrigada por me dares asas e me ensinares a voar bem alto.

Ao meu pequeno grande irmão João Afonso, obrigada por todo o apoio, carinho e amor incondicional. Obrigada por esse abraço aconchegante, é o melhor do mundo. És o meu pequeno Herói. Sinto um grande orgulho em ser tua irmã.

E por fim, dedico todo este trabalho à maior e mais brilhante estrela: o meu Pai.





## Resumo

O mel produzido por abelhas da espécie *Apis mellifera* é um produto de elevado valor, não só económico como também nutricional, com diversos e benefícios para a saúde humana.

De acordo com a sua composição, o mel pode ser classificado em dois tipos:

- monofloral: quando 45 % do seu pólen provém apenas uma espécie floral;
- multifloral: quando provém de mais do que uma espécie floral.

A qualidade do mel varia de acordo como a composição floral, condições climáticas e ambientais da sua zona de produção. É de extrema importância identificar as espécies florais que compõem o mel. A correta identificação de espécies que o constituem, para além de ter impacto no seu valor comercial, assegura a sua autenticidade.

O método clássico usado para identificar as espécies botânicas presentes no mel através da análise dos grãos de pólen é denominado melissopalinologia. Este método baseia-se na observação, identificação e contagem dos grãos de pólen através de microscopia ótica, um método muito moroso, que requer muita experiência e vasto conhecimento acerca da morfologia dos diferentes grãos de pólen.

Nos últimos anos têm sido propostos métodos alternativos à melissopalinologia para a identificação das espécies botânicas nas quais o mel teve origem. Um exemplo são as técnicas de sequenciação de DNA, uma vez que o pólen presente no mel pode servir como fonte de DNA e ser usado para a identificação da sua origem botânica.

O presente trabalho teve como objetivo caracterizar 41 amostras de mel usando uma técnica de *Next-Generation Sequencing* (NGS). A aplicação da técnica de NGS a estas amostras permitiu identificar, ao nível da família, a sua origem botânica. Também foi possível verificar que existe uma relação entre as famílias de plantas identificadas e a origem geográfica do mel. Embora ainda existam algumas limitações relacionadas com a aplicação desta técnica na identificação da origem botânica do mel, num futuro próximo, esta poderá vir a ser aplicada de modo a garantir de forma fiável a autenticidade do produto.

## Palavras chave

*Next-Generation Sequencing*, mel, origem botânica, autenticidade, origem geográfica



## Abstract

Honey produced by *Apis mellifera* bees, is a product with high economic value as well as nutritional with several benefits associated with human health and well-being.

Regarding its classification, honey can be classified in two types:

- monofloral – when 45% of its pollen originates from a single floral species;
- multifloral – when it comes from more than one floral species.

Honey's quality is related with floral composition, climatic and environmental conditions of its production area. It's critical to properly identify the botanic species present in honey once this identification will have an impact on its commercial, also ensuring its authenticity.

The classic method used to identify botanic species within honey is called melissopalynology. This methodology is based on the observation, identification and counting of the pollen grains present in honey by using optical microscopy. This is a very laborious method which requires a high level of expertise and knowledge regarding pollen grains' morphology.

In the last years, alternative methods to melissopalynology that enable the identification of the botanic species within the honey, such as DNA sequencing have been proposed.

The pollen present in the honey can be used as a source of DNA and, therefore used to identify its botanical origin.

This work aimed to characterize 41 honey samples by using Next-Generation Sequencing techniques. This technique allowed the identification of the botanical origin of these honey samples to a family level. It was also possible to establish a link between the identified plants present in the honey samples with their geographic origin. Although there are some constraints related to the application of the NGS technique, in a near future, this technique might be applied to ensure the authenticity of honey.

## Keywords

*Next-Generation Sequencing*, honey, botanical origin, authenticity, geographical origin



# Índice

Agradecimentos .....	v
Resumo .....	vii
Palavras chave.....	vii
Abstract .....	ix
Keywords.....	ix
Lista de Tabelas .....	xii
Lista de Figuras .....	xii
Lista de Abreviaturas .....	xiii
<b>1. Introdução .....</b>	<b>1</b>
1.1. Mel.....	1
1.2. Origem e Características do Mel.....	1
1.3. Classificação do Mel .....	2
1.3.1. Tipos e Representatividade dos Grãos de Pólen .....	2
1.3.2. Frequência dos Grãos de Pólen .....	3
1.4. Autenticidade .....	3
1.5. Melissopalínologia .....	4
1.6. Metodologias Alternativas .....	5
1.6.1. <i>Next Generation Sequencing</i> (NGS) .....	6
<b>2. Objetivo e Estratégia Experimental.....</b>	<b>8</b>
<b>3. Materiais e Métodos .....</b>	<b>11</b>
3.1. Amostras de Mel .....	11
3.2. Escolha do pré-tratamento .....	11
3.3. Escolha do método de extração de DNA.....	13
3.4. Amplificação de DNA por PCR.....	13
3.5. Sequenciação das Amostras de Mel .....	13
3.6. Análise dos Resultados da Sequenciação .....	14
3.6.1. Determinação da Origem Botânica .....	14
3.6.2. Análise de Componentes Principais e Construção do Dendrograma .....	14
3.7. Extração de DNA, Amplificação e Sequenciação de <i>Lavandula stoechas</i> .....	14
<b>4. Resultados e Discussão .....</b>	<b>15</b>
4.1. Escolha do pré-tratamento .....	15
4.2. Escolha do método de extração de DNA.....	19
4.3. Identificação da Origem Botânica.....	21
4.4. Origem Geográfica .....	25
<b>5. Conclusão e Perspetivas Futuras.....</b>	<b>30</b>
<b>6. Bibliografia.....</b>	<b>31</b>
<b>A.Anexos.....</b>	<b>35</b>

## Lista de Tabelas

**Tabela 1.1** Representatividade de alguns tipos de pólen no mel.

**Tabela 3.1** Origem geográfica e classificação das amostras de mel de acordo com os seus produtores.

**Tabela 4.1** Reprodutibilidade da técnica de NGS.

**Tabela A.1** Número médio de sequências obtidas e de famílias de plantas identificadas em 41 amostras de mel pela técnica de NGS.

**Tabela A.2.** Famílias dos grãos de pólen predominantes e secundários identificados através da melissopalinologia e NGS.

## Lista de Figuras

**Figura 1.1** Visualização de grãos de pólen presentes em diferentes amostras de mel através de microscopia ótica.

**Figura 1.2** Fluxograma da técnica de NGS com a plataforma Ion Torrent <sup>TM</sup>.

**Figura 1.3** Imagem ilustrativa do PCR de emulsão.

**Figura 2.1** Fluxograma do delineamento experimental.

**Figura 3.1** Visualização dos grãos de pólen corados com safranina ao microscópio ótico.

**Figura 4.1** Imagem representativa dos grãos de pólen da mesma amostra visualizados ao microscópio ótico após incubação à temperatura de: a) 40 °C e b) 65 °C.

**Figura 4.2** Observação da lise dos grãos de pólen da mesma amostra de mel ao microscópio ótico.

**Figura 4.3** Amplificação de fragmentos de DNA da região ribossomal das amostras submetidas aos diferentes pré-tratamentos.

**Figura 4.4** Amplificação dos produtos de PCR da região ribossomal das amostras extraídas com os diferentes kits de extração de DNA.

**Figura 4.5** Comparação das espécies botânicas identificadas na mesma amostra de mel por melissopalinologia (a) e por NGS (b).

**Figura 4.6** Gráfico de componentes principais com as diferentes amostras de mel.

**Figura 4.7** Dendrograma baseado na distância euclidiana.

**Figura 4.8** Origem geográfica das amostras de mel.

**Figura 4.9** Espécies florais identificadas por NGS de uma amostra de mel oriunda de uma zona de clima temperado mediterrânico.

**Figura 4.10** Espécies florais identificadas por NGS de uma amostra de mel proveniente de uma zona de clima temperado mediterrânico com influência atlântica.

**Figura A.1** Resultado do alinhamento de uma sequência nucleotídica de uma região plastidial de *Erica umbellata* com a ferramenta BLAST.

**Figura A.2** Imagem ilustrativa de famílias de plantas de uma amostra de mel que não foram identificadas pela melissopalinologia (evidenciadas por uma seta).

**Figura A.3** Imagem representativa do resultado do alinhamento das sequências nucleotídicas com a ferramenta BLAST e da falta de poder de resolução dos primers utilizados na identificação da espécie *Eucalyptus sp.*

## Lista de Abreviaturas

<b>BLAST</b>	<i>Basic Local Alignment Tool</i>
<b>CBOL</b>	<i>Consortium for the Barcode of Life</i>
<b>CTAB</b>	Brometo de hexadeciltrimetilamónio
<b>DNA</b>	<i>Deoxyribonucleic acid</i> (Ácido desoxirribonucleico)
<b>MO</b>	Microscópio ótico
<b>NCBI</b>	<i>National Center for Biotechnology Information</i>
<b>NGS</b>	<i>Next-Generation Sequencing</i>
<b>PCA</b>	<i>Principal Component Analysis</i> (Análise de Componentes Principais)
<b>PCR</b>	<i>Polymerase Chain Reaction</i>
<b>PGM</b>	<i>Personal Genome Machine</i>

# 1. Introdução

## 1.1. Mel

De acordo com a União Europeia, mel é a substância natural produzida por abelhas a partir do néctar de plantas. O mel também pode ser produzido através de secreções de diversas plantas, assim como de excreções de afídios nas plantas posteriormente recolhidas pelas abelhas <sup>1,2</sup>.

As abelhas referentes à espécie *Apis mellifera*, são as principais responsáveis pela produção de mel. As abelhas obreiras são as responsáveis pela tarefa de polinizar, recolher e armazenar o néctar das flores nos favos das colmeias. Após a sua recolha, o néctar começa a ser de imediato transformado em mel. Durante este processo de transformação, enzimas existentes nas glândulas salivares das abelhas, nomeadamente uma invertase e uma diástase, hidrolisam açúcares complexos como a sacarose em monossacarídeos como a frutose e a glucose. No final deste processo, o néctar é depositado nos favos da colmeia cuja “arquitetura” estrutural facilita, conjuntamente com o bater das asas das abelhas, a redução em teor de água do néctar (16–20%) dando, assim, origem ao produto que conhecemos como mel <sup>3</sup>.

## 1.2. Origem e Características do Mel

As características do mel variam sobretudo de acordo com o tipo de planta visitada pelas abelhas no momento da recolha do néctar, ou seja, com a sua origem botânica. Mas também, naturalmente, em função da origem geográfica, do clima, das condições de processamento e armazenamento. Todavia, para que possa ser designado como mel, o produto em questão tem de ser constituído essencialmente por açúcares, tais como a frutose e a glucose, cujo teor pode variar entre 30–45% e 24–40%, respetivamente. Para além destes, é possível encontrar no mel açúcares como a trealose, maltose, entre outros. Igualmente presentes estão no mel compostos como polifenóis, ácidos orgânicos, minerais, vitaminas, lípidos e aminoácidos <sup>1,3–6</sup>. O paladar distinto de cada tipo de mel deve-se em grande parte às diferenças na presença deste tipo de compostos <sup>7</sup>.



### 1.3. Classificação do Mel

A classificação do mel é um parâmetro de elevada importância. A determinação e certificação rigorosas da origem botânica e geográfica do mel permitem a sua classificação. Assim, tendo em conta a sua origem botânica, o mel pode ser classificado como:

- Monofloral: quando 45% do seu pólen provém de apenas uma espécie floral;
- Multifloral: quando provém de mais do que uma espécie.

Por exemplo, considera-se que um mel é de urze se mais de 45% dos grãos de pólen presentes pertencerem à planta *Erica sp.* .

A origem geográfica do mel é determinada pela presença de grãos de pólen característicos de cada região.

Para a classificação do mel é necessário ter em conta <sup>8</sup>:

- a) o tipo de grãos de pólen e a sua representatividade;
- b) a sua frequência;

#### 1.3.1. Tipos e Representatividade dos Grãos de Pólen

Os grãos de pólen podem ser nectaríferos ou poliníferos. Considera-se que os grãos de pólen são nectaríferos quando a planta ou flor correspondente produz néctar. Quando a planta não produz néctar, estes são considerados poliníferos. As diferentes representatividades de grãos de pólen devem-se às diferentes taxas de produção de pólen entre diferentes espécies de plantas, temperatura, pluviosidade, ação humana entre outros.

Relativamente à representatividade, é possível categorizar os grãos de pólen em dois tipos: sobre-representados e sub-representados (Tabela 1.1) <sup>8-10</sup>.

Os grãos de pólen sobre-representados são aqueles que apresentam uma percentagem no mel superior à percentagem de contribuição dessa espécie botânica no néctar. Por exemplo, considera-se que um mel é de eucalipto se este apresentar mais de 70% de grãos de pólen de *Eucalyptus sp.* , isto porque o eucalipto produz quantidades elevadas de pólen e estas irão estar sobre-representadas no mel.

No caso dos grãos de pólen sub-representados, a sua representatividade no mel é inferior à percentagem de contribuição dessa mesma planta no néctar. Isto é, as plantas responsáveis pela produção deste tipo de grãos de pólen produzem-nos em baixas quantidades e portanto, a sua percentagem no está sub-representada, como é o caso do mel de rosmaninho (*Lavandula stoechas*).

### 1.3.2. Frequência dos Grãos de Pólen

A classificação do mel em mono ou multifloral é feita com base na frequência relativa dos grãos de pólen presentes. O mel é considerado monofloral quando a frequência relativa de determinada espécie de grãos de pólen é superior a 45% <sup>8,10</sup>. Os grãos de pólen podem ser classificados como predominantes, secundários ou minoritários. Quando a sua frequência é superior a 45% são considerados predominantes. Quando a frequência está no intervalo de 16 e 45%, o tipo de grão de pólen é designado secundário, sendo considerado minoritário se a sua frequência for inferior a 16%.

**Tabela 1.1** Representatividade de alguns tipos de pólen no mel. (Adaptado de Louveaux et al, 1978 e Von Der Ohe et al, 2004)

<b>Pólenes com elevada representatividade</b>	<b>Representatividade normal</b>	<b>Pólenes com baixa representatividade</b>
<i>Eucalyptus sp.</i> (> 70%)	<i>Erica sp.</i> (> 45%)	<i>Citrus sp.</i> (> 10-20%)
<i>Castanea sp.</i> (> 70-90%)	<i>Calluna sp.</i> (> 10-70%)	<i>Carduus sp.</i> (> 5-25%)
<i>Brassica napus</i> (> 60%)	<i>Rosmarinus sp.</i> (> 10-60%)	<i>Lavandula sp.</i> (> 10-40%)
	<i>Thymus sp.</i> (> 10-68%)	<i>Medicago sp.</i> (> 20-30%)
		<i>Taraxacum sp.</i> (> 5-40%)

### 1.4. Autenticidade

O mel é um produto de elevado valor económico, com um vasto potencial de aplicação em várias áreas que vão da cosmética à alimentação. É um produto com um reconhecido valor nutricional, apresentando diversos benefícios para a saúde humana graças às suas propriedades anti-microbianas, anti-virais, anti-inflamatórias, anti-tumorais e antioxidantes <sup>11-13</sup>. As suas propriedades antioxidantes devem-se à presença de compostos fenólicos que atuam nos radicais livres. Para além destas propriedades, diversos estudos associam o consumo de mel à proteção da função cardíaca, hepática e dermatológica, bem como ao controlo de açúcares no sangue e da tensão arterial <sup>14,15</sup>.

O crescente interesse da população em produtos alimentares relacionados com o bem-estar e saúde, aumenta o seu valor económico, tornando-os mais vulneráveis e suscetíveis a fraudes <sup>16,17</sup>.

Segundo o relatório de 2013 do Parlamento Europeu sobre crise alimentar, fraudes na cadeia alimentar e respetivo controlo, a expressão “fraude alimentar” constitui o termo coletivo para designar a substituição, adição, adulteração ou deturpação intencionais dos

gêneros alimentícios, de ingredientes alimentares ou da respetiva embalagem ou informações falsas ou enganadoras prestadas sobre um produto para fins lucrativos <sup>18-20</sup>.

O mel monofloral, apresentando um valor comercial mais elevado comparativamente ao mel multifloral, torna-se particularmente apetecível e sujeito a fraudes, com consequências sérias e imprevisíveis a que importa estar atento.

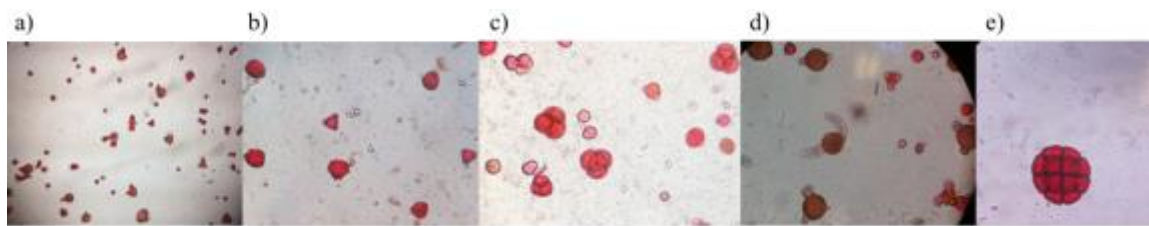
A necessidade de controlar a qualidade e autenticidade de produtos alimentares surge, de certa forma, após a descoberta de uma série de eventos fraudulentos, um pouco por todo o mundo, que puseram em causa a saúde pública. Casos como o da fraude de adulteração de leite na China, em 2008, que levou à morte de 6 bebés <sup>21,22</sup> e o escândalo da carne de cavalo em 2012 <sup>23</sup> tiveram graves impactos não só a nível de saúde pública, como conduziram também a uma diminuição da confiança dos consumidores.

Por essa razão, a União Europeia, através de normas diretivas passou a regulamentar qualidade e autenticidade dos produtos alimentares. Nesse relatório do Parlamento Europeu, o mel foi considerado precisamente um dos dez produtos alimentares mais suscetível de fraude <sup>18</sup> e por isso, tornando mais urgente e necessário certificar a sua origem e qualidade, salvaguardando desta forma a confiança dos consumidores neste produto.

Atualmente, a autenticidade e a qualidade do mel encontram-se regulamentadas através de normas europeias como a *European Honey Directive 2001/110/EC* <sup>2</sup> e o *Codex Alimentarius Standard for Honey* <sup>1</sup>. Assim, a qualidade do mel e a sua classificação são fatores essenciais na valorização e garantia da sua autenticidade que têm cada vez mais de ser respeitadas. Por essa razão, é hoje essencial e urgente o desenvolvimento ou aperfeiçoamento de técnicas que visem garantir a autenticidade e qualidade de produtos alimentares.

### 1.5. Melissopalinologia

O método clássico usado para determinar as espécies botânicas presentes no mel é designado melissopalinologia <sup>10</sup>. Este método consiste na análise dos grãos de pólen presentes no mel através da sua observação, identificação e contagem utilizando microscopia ótica (Figura 1.1).



**Figura 1.1** Visualização de grãos de pólen presentes em diferentes amostras de mel através de microscopia ótica. Os grãos de pólen foram corados com safranina. a) Ampliação 100x; b)–d) ampliação 400x; e) ampliação de 1000x com óleo de imersão.

Como facilmente se depreende da descrição, a melissopalínologia é um método moroso. Para além disso, esta técnica requer um operador especializado e experiente, com elevado conhecimento dos diferentes grãos de pólen, tarefa complicada tendo em conta que, em alguns casos, os grãos de pólen podem apresentar morfologias muito idênticas dificultando assim a sua identificação <sup>8,24,25</sup>. Apesar desta técnica ter as suas desvantagens, continua ainda hoje a ser considerada o *gold-standard* relativamente à identificação da origem botânica do mel.

#### 1.6. Metodologias Alternativas

Nos últimos anos, têm sido propostos outros métodos de determinação da origem botânica do mel alternativos à melissopalínologia, como é o caso das técnicas de sequenciação de DNA, com especial relevância para o uso de técnicas de *Next-Generation Sequencing* (NGS) <sup>24–29</sup>.

Com efeito, os avanços da tecnologia na área da sequenciação de DNA permitiram o desenvolvimento de plataformas que sequenciam milhões de moléculas de DNA em paralelo e a baixo custo <sup>30–33</sup>. Estas técnicas mostram-se vantajosas em relação à melissopalínologia, não apenas por possibilitarem obter resultados precisos e de confiança num curto espaço de tempo, como também permitem a análise simultânea de várias amostras ou de amostras contendo mistura de espécies, como é o caso do mel.

O uso de técnicas de NGS para caracterização de amostras de mel apresenta ainda outras vantagens: estas não requerem um especialista em identificação microscópica de grãos de pólen e, num futuro próximo, poderão permitir uma maior e mais rigorosa discriminação ao nível das espécies botânicas presentes na amostra. Apesar das vantagens referidas, convém acrescentar que, todavia, esta técnica não permite uma análise quantitativa e os grãos de pólen presentes em menores quantidades podem não ser detetados <sup>25,27</sup>.

Dado o carácter recente desta proposta, ainda existem poucos estudos que utilizem diretamente os grãos de pólen presentes no mel para determinação da sua origem botânica. No entanto, estudos recentes demonstraram que o uso de marcadores nucleares e cloroplastidiais como o *ITS2*, o *rbcL* e o *matK* podem ser usados para identificação de pólen através de técnicas de sequenciação<sup>27-29,34,35</sup>.

#### 1.6.1. Next Generation Sequencing (NGS)

Para este trabalho foram sequenciadas diversas amostras de mel com a plataforma Ion Torrent™ da Life Technologies. Esta plataforma de sequenciação de DNA baseia-se no princípio sequenciação por iões semi-condutores. Este tipo de sequenciação consiste na deteção dos iões de hidrogénio que são libertados, provocando a alteração do pH, aquando da incorporação de um determinado nucleótido.

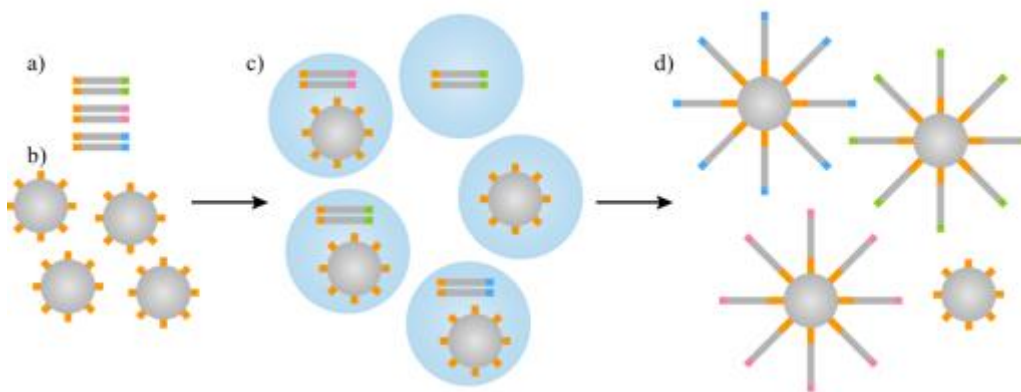
Esta técnica divide-se em quatro fases distintas: a) construção da biblioteca; b) preparação do DNA a ser sequenciado; c) sequenciação e d) análise de dados (Figura 1.2).



**Figura 1.2** Fluxograma da técnica de NGS com a plataforma Ion Torrent™. (Adaptado de [www.thermofisher.com](http://www.thermofisher.com)).

A construção da biblioteca consiste na preparação das moléculas de DNA que irão ser sequenciadas. Este passo consiste na amplificação por PCR das regiões-alvo de estudo das moléculas de DNA com *barcodes*. Para que as amostras possam ser identificadas, é necessário que os produtos de PCR de cada amostra tenham ligado a si um fragmento que permite a identificação de cada uma das amostras utilizadas – os *barcodes*. Os *barcodes* são pequenas sequências conhecidas de nucleótidos que se ligam às moléculas de DNA de cada amostra a ser sequenciada. Após a sequenciação e uma vez que a sequência do *barcode* é conhecida, é possível separar bioinformaticamente as sequências obtidas de cada amostra consoante o *barcode* que foi utilizado. Posteriormente, o produto resultante da amplificação por PCR é purificado e quantificado.

A preparação do DNA a ser sequenciado é feita em duas fases: PCR de emulsão e enriquecimento. O PCR de emulsão ocorre numa matriz de óleo com água e consiste na amplificação clonal dos fragmentos da biblioteca, previamente preparada, em esferas de agarose (Figura 1.3).



**Figura 1.3** Imagem ilustrativa do PCR de emulsão. a) Moléculas de DNA com *barcodes*; b) Esferas de agarose; c) PCR de emulsão numa matriz de óleo; d) Amplificação clonal dos fragmentos. (Adaptado de [www.atdbio.com](http://www.atdbio.com)).

De seguida, os fragmentos amplificados no PCR de emulsão são enriquecidos. Este passo consiste na eliminação das esferas que não contêm DNA. Para tal, são usadas esferas magnéticas com um revestimento de estreptavidina que se liga ao DNA. É então aplicado um campo magnético e feita a seleção das esferas enriquecidas. Posteriormente, um *chip* que contém milhões de poços é carregado com essas esferas e cada uma destas entra num único poço.

De seguida, iniciam-se os ciclos do *chip* no sequenciador. Ou seja, o *chip* é inundado sequencialmente com cada um dos quatro nucleótidos (A, T, G ou C). Sempre que um dos nucleótidos é incorporado na cadeia de DNA, é libertado um ião de hidrogénio ( $H^+$ ). A libertação de prótons provoca uma alteração no pH da solução do poço. Esta alteração de pH é detetada, medida e convertida em voltagem no sequenciador Ion Personal Genome Machine (PGM™). Este processo ocorre em simultâneo nos milhões de poços de cada *chip*.

Concluído o processo de sequenciação, obtêm-se as sequências das moléculas de DNA em formato FASTAQ. A sua identificação é possibilitada através da comparação das sequências obtidas com as sequências existentes em bases de dados.

## 2. Objetivo e Estratégia Experimental

Este trabalho teve como objetivo caracterizar amostras de mel através de métodos moleculares (*Next-Generation Sequencing*). De forma a avaliar a viabilidade deste método, os resultados obtidos por sequenciação foram comparados com os resultados obtidos pelo método clássico, a melissopalínologia. Assim, pretendeu-se avaliar o uso das técnicas NGS na identificação da origem botânica de amostras de mel de maneira a garantir a autenticidade e qualidade deste produto.

De modo a alcançar o objetivo, o primeiro passo do trabalho foi a realização de uma extensa pesquisa bibliográfica focando pré-tratamentos e métodos de extração de DNA de grãos de pólen presentes no mel. De acordo com a sua constituição, a consistência do mel pode variar entre um estado totalmente líquido ou quase sólido, o que pode dificultar o seu processamento. Para além disso, trata-se de um composto com muitos açúcares e polifenóis que são conhecidos inibidores da reação de PCR<sup>16,25,36,37</sup>. Adicionalmente, o número de grãos de pólen presentes no mel pode ser extremamente variável, sendo este um fator determinante na eficácia da extração do seu DNA. Tendo estes fatores em consideração, tornou-se imperativo o desenvolvimento de um protocolo capaz de contornar estas limitações (Figura 2.1).

Assim sendo, a parte inicial deste projeto focou-se em duas etapas experimentais:

- 1) Escolha de um pré-tratamento inicial – de maneira a facilitar o processamento das amostras de mel e solubilizar os açúcares nele presentes. Para tal, foram testados diferentes parâmetros como:
  - i. Temperatura de incubação – que teve como finalidade ajudar no processamento das amostras devido à sua consistência variável, assim como solubilizar os açúcares presentes no mel. Testaram-se duas temperaturas diferentes: 40 °C e 65 °C.;
  - ii. Tipo de lise – por lise entende-se o processo de rutura da parede celular dos grãos de pólen. Consideraram-se para este trabalho dois tipos de lise: mecânica e mecânica e química.
- 2) Escolha do método de extração de DNA mais adequado para os grãos de pólen – foram testados diferentes kits comerciais de extração de DNA.

Os métodos de extração compreendem essencialmente 3 passos: a lise das células, a purificação e a recuperação do DNA<sup>38</sup>.

Os kits comerciais testados baseiam-se no método de extração por coluna. Isto consiste no uso de uma coluna que contém uma membrana de sílica com afinidade para DNA. A utilização de *binding buffers* com uma concentração elevada de sais promove a ligação do DNA à membrana da coluna. De seguida, as impurezas e os contaminantes presentes são removidos através de um tampão com um detergente e etanol. Por último, o DNA é eluído com um tampão com uma baixa concentração de sais, promovendo a sua libertação da membrana da coluna.

É da maior importância selecionar um método de extração apropriado, uma vez que a quantidade e qualidade de DNA irão influenciar todos os procedimentos seguintes.

Durante a fase de otimização do pré-tratamento e extração do DNA, foi usada sempre a mesma amostra de mel de maneira a introduzir o mínimo de variabilidade possível.

Após a escolha do pré-tratamento e do método de extração de DNA, procedeu-se à amplificação das regiões-alvo de estudo através da reação de *Polymerase Chain Reaction* (PCR). As zonas alvo de amplificação encontram-se universalmente distribuídas pelas plantas e possuem zonas conservadas e zonas variáveis. As regiões conservadas contêm um número suficiente de locais informativos comuns ao grupo das plantas, enquanto que as regiões variáveis diferem entre os diversos grupos, o que possibilita a sua identificação.

Para identificar as espécies botânicas presentes no mel, realizou-se a sequenciação por NGS de 41 amostras.

Paralelamente, e de forma independente, todas as amostras foram submetidas à análise melissopalínológica.



### 1ª Etapa

#### Escolha do pré-tratamento

- Temperatura de incubação
- Lise mecânica ou mecânica e química

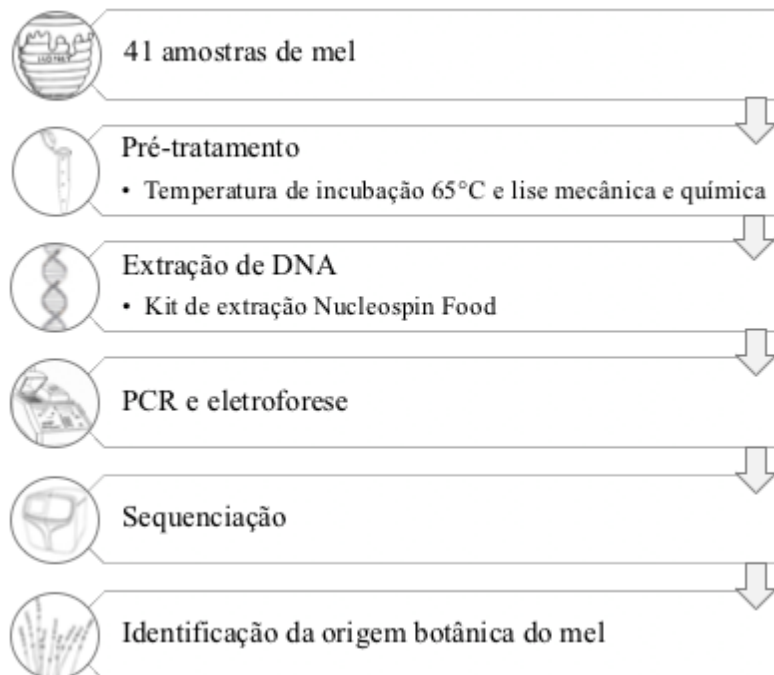


### 2ª Etapa

#### Escolha do método de extração de DNA dos grãos de pólen com os kits:

- DNeasy Plant Mini
- Plant DS
- DNeasy mericon Food
- Tissue DNA
- Nucleospin Food

### Estratégia Completa



**Figura 2.1** Fluxograma do delineamento experimental.

### 3. Materiais e Métodos

#### 3.1. Amostras de Mel

No decorrer deste estudo, foram estudadas 41 amostras de mel fornecidas pelo Instituto de Biologia Experimental e Tecnológica (iBET).

A origem botânica de cada amostra foi indicada pelos seus produtores com base na localização da colmeia e na flora circundante. As amostras estudadas são provenientes de doze distritos de Portugal Continental (Tabela 3.1).

Todas as amostras foram armazenadas num local seco, protegidas da luz e à temperatura ambiente. O seu processamento inicial foi realizado numa câmara de fluxo laminar de forma a prevenir contaminações.

**Tabela 3.1** Origem geográfica e classificação das amostras de mel de acordo com os seus produtores.

<b>Amostra</b>	<b>Distrito</b>	<b>Tipo de Mel</b>	<b>Amostra</b>	<b>Distrito</b>	<b>Tipo de Mel</b>
MEL 1	Lisboa	Multi floral	MEL 34	Braga	Multi floral
MEL 2	Lisboa	Multi floral	MEL 35	Guarda	Carvalho
MEL 5	Beja	Rosmaninho	MEL 36	Santarém	Multi floral
MEL 9	Viseu	Multi floral	MEL 37	Lisboa	Multi floral
MEL 10	Setúbal	Multi floral	MEL 40	Santarém	Urze
MEL 13	Faro	Multi floral	MEL 41	Santarém	Alecrim
MEL 15	Évora	Multi floral	MEL 53	Viseu	Multi floral
MEL 16	Lisboa	Multi floral	MEL 54	Viseu	Multi floral
MEL 17	Bragança	Multi floral	MEL 55	Viseu	Rosmaninho
MEL 18	Bragança	Urze	MEL 64	Beja	Multi floral
MEL 19	Castelo Branco	Urze	MEL 72	Castelo Branco	Azinheira
MEL 20	Santarém	Multi floral	MEL 73	Castelo Branco	Castanheiro
MEL 23	Setúbal	Multi floral	MEL 74	Castelo Branco	Rosmaninho
MEL 24	Santarém	Multi floral	MEL 75	Santarém	Urze
MEL 25	Coruche	Eucalipto	MEL 76	Santarém	Urze
MEL 26	Braga	Multi floral	MEL 77	Castelo Branco	Laranjeira
MEL 28	Braga	Eucalipto	MEL 79	Castelo Branco	Multi floral
MEL 29	Braga	Multi floral	MEL 80	Coimbra	Urze
MEL 30	Braga	Multi floral	MEL 101	Santarém	Urze
MEL 32	Braga	Multi floral	MEL PT1	Castelo Branco	-
MEL 33	Braga	Urze			

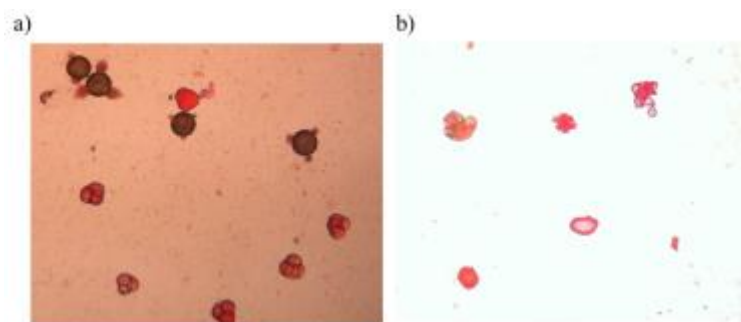
#### 3.2. Escolha do pré-tratamento

As amostras de mel foram diluídas em água numa proporção de 1:4. De seguida foram incubadas num banho termostático (VWB12, VWR, USA) a 40 °C ou a 65 °C durante 30 minutos e centrifugadas durante 20 minutos a 4000 rpm (Eppendorf 5810R, GmbH,

Hamburg, Germany). O *pellet* formado após a centrifugação foi ressuspendido e procedeu-se à lise dos grãos de pólen com o auxílio de um vórtex. Foram testados dois tipos de lise:

- i) Mecânica – tendo sido testadas diferentes tipos de esferas (microesferas de vidro (425-500  $\mu\text{M}$ , Sigma-Aldrich), esferas de vidro com 1.5-2 mm de diâmetro, *Lysing Matrix C*® e *Lysing Matrix E*® (MP Biomedicals, Solon, OH)). As microesferas e as esferas de vidro são compostas por vidro, as *Lysing Matrix C*® são esferas de sílica e *Lysing Matrix E*® são uma mistura de esferas de vidro, sílica e cerâmica.
- ii) Mecânica e química – utilizando as diferentes esferas referidas acima e tampão de extração contendo CTAB (brometo de hexadeciltrimetilamónio –  $(\text{C}_{16}\text{H}_{33})\text{N}(\text{CH}_3)_3\text{Br}$ ). O CTAB ajuda a remoção de polissacarídeos presentes na amostra e a promover a solubilização da parede celular.

A lise dos grãos de pólen foi confirmada por observação com microscópio ótico (Leica, Germany) (Figura 3.1).



**Figura 3.1** Visualização dos grãos de pólen corados com safranina ao microscópio ótico. a) Grãos de pólen antes da lise; b) grãos de pólen após lise mecânica e química com esferas de vidro. Ampliação 100x.

Para avaliar a eficiência dos diferentes pré-tratamentos, procedeu-se à extração do DNA com o kit Nucleospin Food (Macherey-Nagel GmbH, Germany) seguindo as instruções do fornecedor. Uma vez selecionado o pré-tratamento, procedeu-se à otimização da extração do DNA com diferentes kits comerciais.

### 3.3. Escolha do método de extração de DNA

Após a escolha do pré-tratamento, o passo seguinte focou-se na determinação do método de extração de DNA. Foram utilizados os kits comerciais: Plant DS kit e Tissue DNA kit (Omega Bio-tek, GA, USA), DNeasy mericon Food kit e DNeasy Plant Mini kit (Qiagen, Valencia, CA, USA), e Nucleospin Food (Macherey-Nagel GmbH, Germany). As extrações de DNA realizaram-se de acordo com as instruções do fornecedor de cada kit.

Após a extração do DNA, as amostras foram quantificadas com o dsDNA BR Assay Kit do Qubit® fluorometer 2.0 (Life Technologies). Realizou-se também uma eletroforese em gel de 1% de agarose com o DNA extraído para avaliar a sua qualidade e quantidade.

### 3.4. Amplificação de DNA por PCR

Foram amplificadas regiões do DNA nuclear ribossomal e plastidiais com dois pares de *primers* específicos (propriedade da Biopremier). As regiões utilizadas são reconhecidas como tendo potencial discriminatório ao nível do género e, em alguns casos, da espécie. A reação de amplificação ocorreu num termociclador Life ECO (Bioer, China), tendo as condições sido anteriormente otimizadas na Biopremier.

A amplificação dos fragmentos de PCR foi confirmada por eletroforese (gel de 1,5% de agarose com 5 µl dos produtos do PCR, tendo o gel sido corado com brometo de etídio, um agente intercalante de DNA visível com luz ultravioleta). Como marcador de massa molecular, utilizou-se o GeneRuler 100bp Plus DNA Ladder (Thermo Fisher).

Como controlo positivo foi usado o DNA previamente extraído de uma planta da família Fabaceae.

### 3.5. Sequenciação das Amostras de Mel

A sequenciação foi realizada com a plataforma Ion Torrent™ (Life Technologies).

A preparação e purificação das bibliotecas para o sequenciador PGM™ foi feita de acordo com o protocolo previamente otimizado na Biopremier.

Os PCR de emulsão e enriquecimento foram realizados com recurso aos equipamentos *Ion OneTouch 2 System* ou *Ion Chef* e todas as bibliotecas foram sequenciadas em *chips* 314 ou 316.

### 3.6. Análise dos Resultados da Sequenciação

#### 3.6.1. Determinação da Origem Botânica

As sequências obtidas em formato FASTAQ foram extraídas de acordo com o *barcode* e *primers* utilizados e agrupadas utilizando um *software* desenvolvido na Biopremier. Posto isto, procedeu-se à comparação das sequências resultantes da sequenciação com as sequências disponíveis na base de dados pública GenBank (NCBI – <http://www.ncbi.nlm.nih.gov/genbank/>) e ao seu alinhamento com o auxílio da ferramenta *Basic Local Alignment Tool* (BLAST). Isto permitiu a identificação de diversas famílias, e em alguns casos, espécies de plantas presentes nas amostras de mel. Foram selecionadas, de forma aleatória, 10% das amostras para realização de triplicados para avaliar a reprodutibilidade da metodologia utilizada.

#### 3.6.2. Análise de Componentes Principais e Construção do Dendrograma

Através da análise de componentes principais (PCA) avaliou-se a relação entre as famílias botânicas detetadas pela sequenciação em cada amostra de mel com a sua origem geográfica. A PCA transforma um conjunto de variáveis não correlacionadas, neste caso, as famílias das espécies botânicas encontradas nas amostras de mel, em componentes principais num espaço ortogonal. Este tipo de análise visa explicar a máxima variância entre o menor número de componentes principais.

Através do cálculo da distância euclidiana foi construído um dendrograma de distância. Para tal, foi usado o programa NTSYS (versão 2.21L, Applied Biostatistics Inc.)<sup>39</sup>.

### 3.7. Extração de DNA, Amplificação e Sequenciação de *Lavandula stoechas*

Foi extraído o DNA de uma planta de *Lavandula stoechas* (rosmaninho) com o kit Nucleospin Food (Macherey-Nagel GmbH, Germany), através da maceração das suas folhas. Amplificaram-se por reação de PCR as regiões ribossomais e plastidiais aqui estudadas com *primers* específicos. Após a confirmação da amplificação dos produtos de PCR, a Biopremier procedeu à sua sequenciação por Sanger. Com recurso à ferramenta BLAST, os resultados da sequenciação foram comparados e alinhados com as sequências obtidas pela técnica de NGS pertencentes à famílias Lamiaceae.

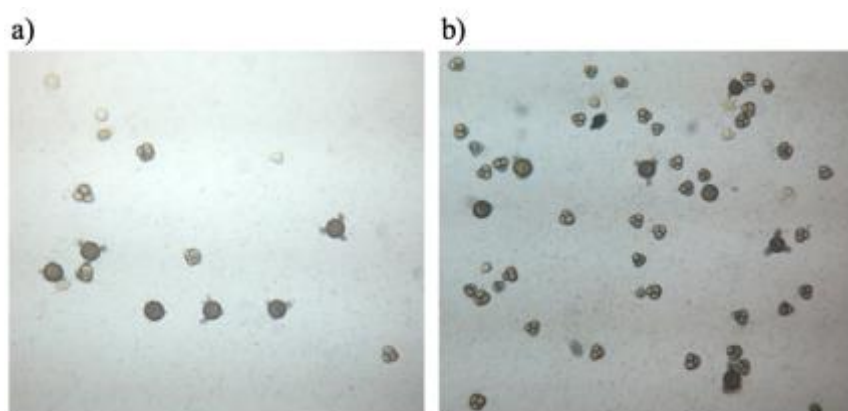
## 4. Resultados e Discussão

### 4.1. Escolha do pré-tratamento

A existência de um pré-tratamento inicial é de extrema importância, uma vez que o mel é um produto de consistência variável e é constituído maioritariamente por açúcares e polifenóis. A presença destes pode ter como consequência a inibição da reação PCR. Assim, torna-se importante a existência de um passo de pré-tratamento das amostras que facilite o seu processamento e maximize a dissolução de inibidores. Para além disso, os grãos de pólen existentes no mel são estruturas celulares extremamente complexas e resistentes, o que dificulta a sua lise.

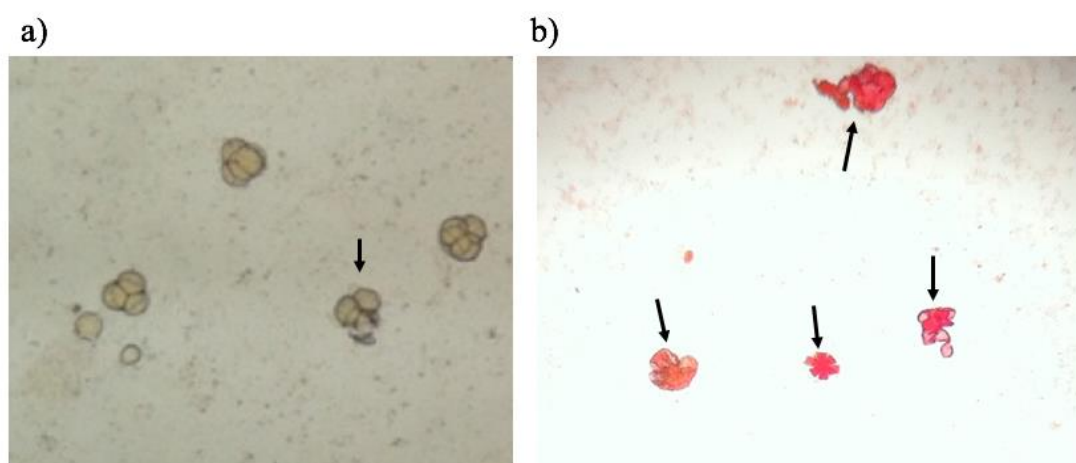
Neste trabalho, o pré-tratamento consistiu em dois passos distintos: incubação das amostras e lise dos grãos de pólen. Primeiro, foram testadas duas temperaturas de incubação das amostras de mel: 40 °C e 65 °C. Após incubação, foi possível observar que o mel das amostras incubadas a 65 °C se encontrava totalmente solubilizado, ao contrário das amostras cuja temperatura de incubação foi de 40 °C. Estas apresentavam alguma heterogeneidade. Para além disto, após centrifugação do mel solubilizado a 65 °C, observou-se uma maior sedimentação de material e de forma mais eficiente do que aquele cujo pré-tratamento incluiu uma temperatura de solubilização de 40 °C.

Estes resultados sugeriram que a 65 °C a recuperação de grãos de pólen é mais elevada, o que foi confirmado através da observação com microscópio ótico (Figura 4.1). É importante salientar que a recuperação final dos grãos de pólen é um fator determinante – teoricamente quanto maior for a recuperação de grãos de pólen, maior será a quantidade de DNA a extrair.



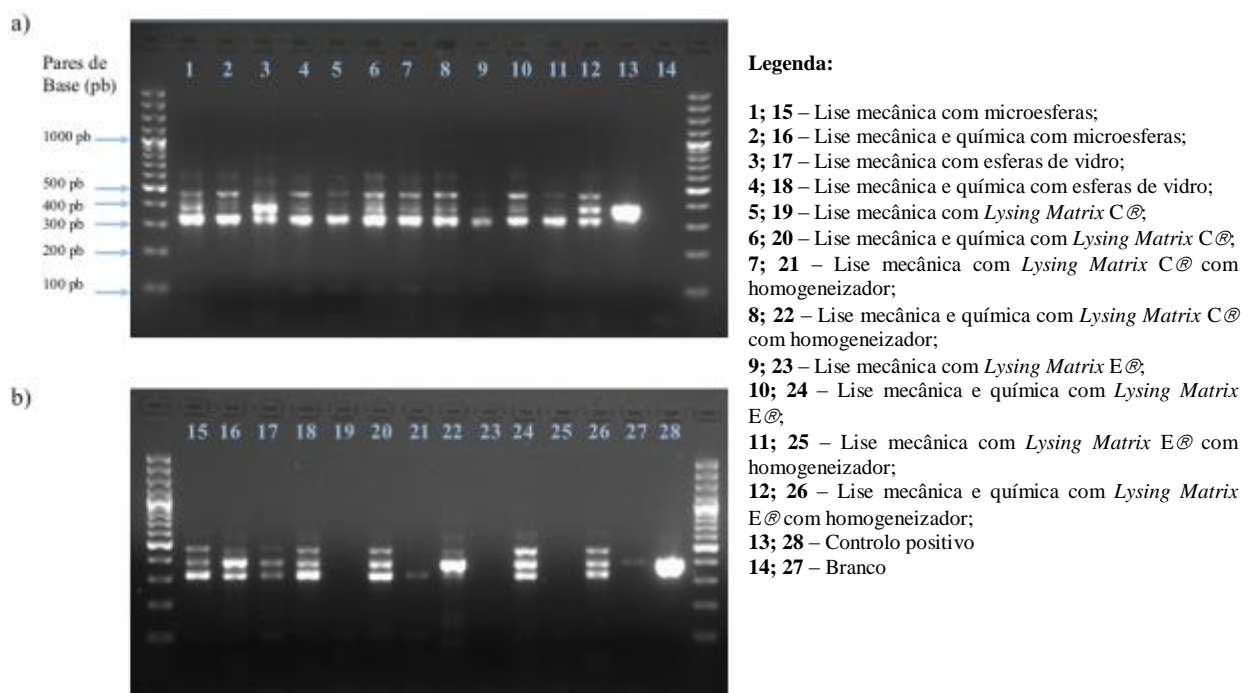
**Figura 4.1** Imagem representativa dos grãos de pólen da mesma amostra visualizados ao microscópio ótico após incubação à temperatura de: a) 40°C e b) 65°C. Ampliação 100x.

Em segundo lugar testaram-se dois tipos de lise celular: mecânica e mecânica e química. Relativamente à lise mecânica, avaliou-se a eficiência de quatro tipos de esferas: as microesferas, as esferas de vidro, as *Lysing Matrix C*® e *E*®. O tipo de lise mecânica e química consistiu no uso das diferentes esferas acima referidas e de um tampão de extração com CTAB. Está descrito na literatura que o CTAB ajuda a eliminar polissacarídeos presentes nas amostras e a promove a solubilização da parede celular<sup>40</sup>. Os resultados obtidos mostraram que a lise mecânica é menos eficaz do que a lise mecânica e química. Isto foi demonstrado através da confirmação da lise dos grãos de pólen com microscópio ótico. A observação ao microscópio ótico mostrou que as amostras lisadas mecanicamente exibiam número inferior considerável de grãos de pólen lisados (Figura 4.2). Também se pode observar que, independentemente da temperatura de incubação das amostras, a lise mecânica foi menos eficaz.



**Figura 4.2** Observação da lise dos grãos de pólen da mesma amostra de mel ao microscópio ótico. As setas evidenciam os grãos de pólen lisados à temperatura de 65 °C submetidas ao tipo de lise a) mecânica e b) mecânica e química. Ampliação de 400x.

Procedeu-se à extração do DNA dos grãos de pólen presentes no mel e à sua amplificação por PCR com um par de *primers* específicos que gerou fragmentos com aproximadamente 340 pb. Este passo permitiu inferir acerca da temperatura de incubação das amostras de mel, confirmar a eficácia da lise celular e avaliar a eficiência de amplificação do DNA. Para tal, realizou-se uma eletroforese em gel de agarose com os produtos de PCR amplificados das amostras que foram submetidas a todos os pré-tratamentos (Figura 4.3).



**Figura 4.3** Amplificação de fragmentos de DNA da região ribossomal das amostras submetidas aos diferentes pré-tratamentos. Temperatura de incubação das amostras: a) 40°C; b) 65°C. Imagem representativa de dois géis de 1,5% de agarose corado com brometo de etídio.

Relativamente à temperatura de incubação das amostras de mel, os fragmentos de DNA visualizados no gel de agarose cuja incubação foi de 40 °C apresentam algum tipo de degradação do DNA. É possível inferir que esta temperatura não é suficiente para remover totalmente açúcares, polifenóis e outros contaminantes na totalidade, dificultando, consequentemente, a lise dos grãos de pólen. Este resultado reforçou os resultados obtidos anteriormente, e permitiu definir a temperatura de incubação (65 °C). Em relação ao tipo de lise, a lise mecânica e química mostrou ser mais eficaz do que a lise mecânica. Os fragmentos de DNA das amostras submetidas à lise mecânica e química apresentaram uma maior eficiência de amplificação. Este resultado permite afirmar que a quantidade e qualidade dos fragmentos de DNA obtidos é mais elevada quando se aplica este tipo de lise. Para além disso, também se observou que houve amplificação de todas as amostras lisadas mecânica e quimicamente enquanto que, em alguns casos, não houve qualquer tipo de amplificação com a lise mecânica. Foram também testados os diferentes tipos de esferas utilizados para a lise mecânica. A eficácia de lise de cada para cada tipo de esferas foi comparada, tendo como base a



observação ao microscópio e amplificação dos produtos de PCR. Constatou-se que as esferas de vidro eram as mais adequadas.

A análise do gel de eletroforese com os produtos de PCR amplificados mostrou que as microesferas são aquelas que apresentam uma amplificação menor. Como já foi referido, os grãos de pólen são estruturas celulares extremamente resistentes e é possível que, devido ao seu pequeno tamanho/diâmetro, este tipo de esferas não consiga lisar na totalidade os grãos de pólen. Embora o resultado da amplificação dos fragmentos de DNA que incluíram a lise com as *Lysing Matrix C*® e *E*® tenha sido positivo, as esferas de vidro foram as selecionadas. Os resultados da amplificação dos produtos PCR das amostras lisadas com as esferas de vidro e com ambas as *Lysing Matrix*® são equiparáveis. No entanto, resultados anteriores, como a confirmação da lise ao microscópio ótico mostraram que as esferas de vidro são mais eficazes. A observação com o microscópio ótico também mostrou que as amostras lisadas com as *Lysing Matrix C*® e *E*® apresentam alguma “sujidade” a rodear os grãos de pólen. Isto pode influenciar a lise, na medida em que os grãos de pólen que se encontram rodeados/protegidos por essa sujidade, podem não ser lisados. Para além disto, as *Lysing Matrix E*® mostraram ser de difícil manuseamento devido ao pequeno tamanho das esferas de sílica presentes nestas matrizes.

Paralelamente, testou-se o uso de um homogeneizador (Precellys, Bertin) para avaliar se o seu uso teria algum efeito na lise dos grãos de pólen. Este foi utilizado nas amostras em que se usaram as *Lysing Matrix C*® e *E*®. Com o resultado da amplificação dos fragmentos de DNA das amostras em que se utilizou o homogeneizador, concluiu-se que o uso deste não influenciava consideravelmente a lise dos grãos de pólen.

Tendo em conta resultados obtidos em termos de recuperação dos grãos de pólen, dissolução de açúcares e polifenóis, eficiência de lise celular e amplificação de DNA, o pré-tratamento foi selecionado o que incluía a incubação das amostras de mel a 65 °C e lise mecânica e química com esferas de vidro. O pré-tratamento escolhido foi aplicado às 41 amostras de mel utilizadas no decorrer do trabalho.

#### 4.2. Escolha do método de extração de DNA

A segunda etapa do trabalho visou determinar o método de extração de DNA que permitisse a obtenção de DNA em maior e melhor qualidade. Testaram-se 5 kits comerciais diferentes cujo princípio assenta no método de extração por coluna.

Após a extração de DNA, todas as amostras foram quantificadas utilizando o Qubit®, no entanto estas mostraram ser não quantificáveis – a concentração do DNA estava abaixo da gama de deteção do Qubit®. Posteriormente, avaliou-se a integridade do DNA genómico de cada amostra de mel através da realização de uma eletroforese em gel de 1% de agarose. No entanto, não foi possível observar DNA genómico para a maioria das amostras. O número de grãos de pólen é altamente variável consoante a amostra de mel estudada. Para além disso, a densidade de grãos de pólen presentes no mel pode ser insuficiente para permitir que o seu DNA seja visível. Ainda assim, procedeu-se à amplificação por PCR do DNA extraído com os diferentes kits utilizando um par *primers* específicos para uma região ribossomal. Isto permitiu avaliar a amplificabilidade do DNA presente, ou não, em cada amostra.

Realizou-se uma eletroforese com os produtos de PCR amplificados comparando-se os resultados obtidos com os diferentes kits de extração (Figura 4.4).

Uma vez que todas as amostras foram amplificadas, foi possível afirmar que ocorreu o correto emparelhamento dos *primers* utilizados.

Através da análise do gel de eletroforese com os produtos de PCR amplificados verificou-se que o kit Nucleospin Food foi aquele que apresentou melhores resultados. O kit Nucleospin Food está otimizado para extração de DNA de amostras alimentares<sup>41</sup>. Comparativamente aos outros kits de extração, a amplificação dos produtos de PCR das amostras extraídas com o kit Nucleospin Food foi a que demonstrou ter uma maior eficiência. Isto traduz-se numa maior quantidade e qualidade dos produtos de PCR obtidos.

Relativamente aos outros kits testados, a amplificação dos fragmentos de DNA extraídos com os kits Tissue DNA e DNeasy mericon Food foi a que demonstrou ser menos eficiente.



**Figura 4.4** Amplificação dos produtos de PCR da região ribossomal das amostras extraídas com os diferentes kits de extração de DNA. Figura representativa de um gel de 1,5% de agarose corado com brometo de etídio.

O kit Tissue DNA está otimizado para a extração de DNA de tecidos e os grãos de pólen, ao contrário de tecidos animais, são mais complexos e difíceis de lisar devido à sua espessa parede celular externa constituída por exina e esporopolenina <sup>42</sup>. É possível afirmar que o kit Tissue DNA não é mais adequado para a extração de DNA de grãos de pólen presentes no mel, uma vez que a amplificação dos fragmentos de DNA foi pouco eficiente. O kit DNeasy mericon Food, embora desenvolvido para extração de DNA de produtos alimentares, utiliza clorofórmio – um composto desnaturante de proteínas. Este composto, que promove a separação física das proteínas do DNA, forma uma interface fenólica-aquosa, sendo a aquosa aquela onde permanece o DNA <sup>43</sup>. É possível inferir que, aquando da formação da interface fenólica-aquosa, a separação do DNA, das proteínas e dos restantes compostos não tenha sido suficiente, provocando consequentemente uma amplificação dos produtos de PCR menos eficaz.

Os kits Plant DS e DNeasy Plant Mini são especialmente indicados para a extração de DNA de tecidos vegetais <sup>44,45</sup>. Ambos demonstraram ser eficazes na extração do DNA das amostras, pois observou-se amplificação dos fragmentos de DNA extraídos com estes kits.

Uma vez que com kit Nucleospin Food se obtiveram os melhores resultados em termos de amplificação de produtos de PCR, este foi o selecionado para a extração de DNA das restantes amostras de mel.

#### 4.3. Identificação da Origem Botânica

Assim que o pré-tratamento e o método de extração foram definidos, procedeu-se à análise das 41 amostras de mel. Todas elas foram submetidas ao pré-tratamento inicial e o seu DNA foi extraído.

O passo seguinte foi a amplificação das regiões-alvo (ribossomal e plastidial) por PCR. As regiões-alvo amplificadas apresentam regiões conservadas que permitem a identificação ao nível do género, e em alguns casos, da espécie botânica. Todas as amostras cuja amplificação destas regiões foi bem sucedida, foram posteriormente sequenciadas com a plataforma Ion Torrent <sup>TM</sup>. Sequenciaram-se por NGS, 41 amostras de mel (Tabela A.1). Após a sequenciação, as sequências obtidas foram identificadas através da sua comparação com sequências disponíveis na base de dados GenBank e alinhamento utilizando a ferramenta BLAST (Figura A.1).

Paralelamente, as 41 amostras de mel foram submetidas à análise melissopalinológica. Esta análise permitiu a identificação da origem botânica das amostras de mel – foram identificados os géneros e/ou espécies dos grãos de pólen predominantes, secundários e minoritários.

Foram comparadas as espécies botânicas identificadas por ambas as técnicas. Após a comparação dos resultados obtidos, constatou-se que o uso da técnica de NGS e de *primers* específicos permitiu:

- a) a identificação da maioria das famílias de plantas dos grãos de pólen nectaríferos predominantes, no caso de existirem, e secundários presentes nas 41 amostras de mel (Figura 4.5 e Tabela A.2);
- b) a identificação de famílias de plantas que não foram identificadas pela melissopalinologia (Figura A.2).

a)

Classe frequencia / Frequency class	Tipo Polínico / Pollinic type	%
Pólen predominante / Predominant pollen	<i>Erica spp.</i>	49
Pólen secundário / Secondary pollen		
Pólen minoritário / Minority pollen	<i>Cytisus / Ulex</i>	10
	<i>Rubus spp. / Eriobotrya spp.</i>	7
	<i>Eucalyptus spp.</i>	6
	<i>Lavandula stoechas</i>	5
	<i>Echium plantagineum</i>	5
	<i>Rhamnus spp.</i>	5
	<i>Salix spp.</i>	4

b)

Família	Espécie
Ericaceae	<i>Erica sp.</i>
Rosaceae	<i>Rubus sp.</i>
Fabaceae	<i>Cytisus striatus</i>
Myrtaceae	<i>Myrtus communis</i>
Boraginaceae	<i>Echium sp.</i>
	<i>Borago officinalis</i>

**Figura 4.5** Comparação das espécies botânicas identificadas na mesma amostra de mel por melissopalínologia (a) e por NGS (b).

Os resultados aqui obtidos são extremamente vantajosos sob o ponto de vista de autenticidade alimentar. Esta técnica mostrou possível a identificação das espécies botânicas que constituem o mel. Este é um ponto de extrema importância porque através desta identificação é possível assegurar a sua autenticidade e identificar possíveis fraudes.

Em alguns casos conseguiu-se fazer a identificação do género, não sendo possível determinar a espécie botânica. Os resultados obtidos mostraram que os *primers* utilizados apresentam alguma falta de poder discriminatório. Estes não permitem fazer a distinção entre espécies pertencentes à família Brassicaceae e espécies do género *Quercus* e *Eucalyptus* (Figura A.3).

O uso da técnica de NGS na identificação das plantas envolvidas na produção de mel, em alguns casos, não permitiu a identificação das mesmas famílias identificadas pela melissopalínologia podendo dever-se a fatores como:

- i. no caso de a lise dos grãos de pólen não ter sido eficiente, o seu DNA não ter sido extraído, não sendo, consequentemente, sequenciado;

- ii. consoante as amostras de mel, os grãos de pólen presentes apresentarem representatividades diferentes e aqueles que são sub-representados poderão não ser detetados na sequenciação;
- iii. a falta de bases de dados públicos que contemplem sequências de DNA das regiões-alvo.

Como já foi mencionado acima, os grãos de pólen são estruturas resistentes à lise e consoante a planta que o produz, uns grãos podem ser mais facilmente lisados do que outros o que influenciará a sua sequenciação. Mais, o mel pode apresentar uma grande diversidade em termos de grãos de pólen presentes e estes podem estar mais ou menos representados. Para a determinação da origem botânica do mel, a melissopalínologia apenas contempla os grãos de pólen nectaríferos, sendo excluídos os grãos de pólen poliníferos <sup>8,10</sup>. Diferentemente, na sequenciação, o DNA presente na amostra, seja este proveniente de grãos de pólen nectaríferos ou poliníferos, irá ser sequenciado. Se a quantidade de DNA de grãos de pólen nectaríferos for inferior à quantidade de DNA de grãos poliníferos, estes poderão ser sequenciados preferencialmente, o que pode resultar na falta de deteção e identificação de determinados grãos de pólen. Para além destas razões, o mel não contém única e exclusivamente DNA de origem vegetal: a presença de fungos neste tipo de produtos está igualmente bem descrita na literatura <sup>46–48</sup>. No decorrer deste estudo, um dos dois pares de *primers* utilizado apresenta a capacidade de amplificar DNA de eucariotas, e consequentemente, DNA de fungos. Logo, o DNA extraído, posteriormente amplificado e sequenciado com este par de *primers* pode permitir, para além da identificação das famílias botânicas que constituem o mel, a identificação de fungos aí presentes. Relativamente à falta de sequências em bases de dados públicas, esta lacuna pode dificultar a identificação das família e/ou espécies que constituem o mel. Após comparação dos resultados da sequenciação por NGS com os resultados da melissopalínologia, observou-se que a base de dados utilizada não contém a sequência da região amplificada do rosmaninho (*Lavandula stoechas*). Isto foi comprovado quando se procedeu à extração e sequenciação das mesmas regiões-alvo de uma planta de rosmaninho. Após a sequenciação, foi feita a comparação e alinhamento da sequência obtida com as sequências disponíveis no GenBank e constatou-se que esta base de dados não contempla o rosmaninho. Até à data, a base de dados utilizada, não

dispõe sequências das regiões-alvo de *Lavandula stoechas*, que pertence à família Lamiaceae. No entanto, é importante salientar que a família (Lamiaceae) é detetada. Este tipo de lacuna dificulta a identificação de determinadas espécies presentes no mel, sendo por isso de extrema importância contornar esta limitação através da atualização das bases de dados ou da criação novas que contemplem sequências de DNA das plantas conhecidas cujo néctar é utilizado na produção de mel.

Mais, a reprodutibilidade desta técnica de sequenciação foi avaliada através da realização de triplicados de 10% das amostras. A sequenciação de DNA através de NGS provou ser extremamente reprodutível. Para cada amostra que se realizou o triplicado da sequenciação, os resultados da identificação, assim como a percentagem das sequências detetadas, estiveram sempre em concordância (Tabela 4.1).

**Tabela 4.1** Reprodutibilidade da técnica de NGS.

Sequenciação 1	Familia	Espécie	% Sequências
		<i>Trigonella cretica</i>	35,87
Fabaceae		<i>Ulex sp.</i>	17,96
		<i>Medicago sp.</i>	12,82
	Rosaceae	<i>Prunus sp.</i>	15,17
Rhamnaceae		<i>Rhamnus sp.</i>	5,72
Oxalidaceae		<i>Oxalis pes-caprae</i>	3,77
Myrtaceae		<i>Eucalyptus sp.</i>	2,79

Sequenciação 2	Familia	Espécie	% Sequências
		<i>Trigonella cretica</i>	35,87
Fabaceae		<i>Ulex sp.</i>	20,02
		<i>Medicago sp.</i>	7,73
	Rosaceae	<i>Prunus sp.</i>	19,70
Rhamnaceae		<i>Rhamnus sp.</i>	9,23
Oxalidaceae		<i>Oxalis sp.</i>	5,17
Myrtaceae		<i>Eucalyptus sp.</i>	2,28

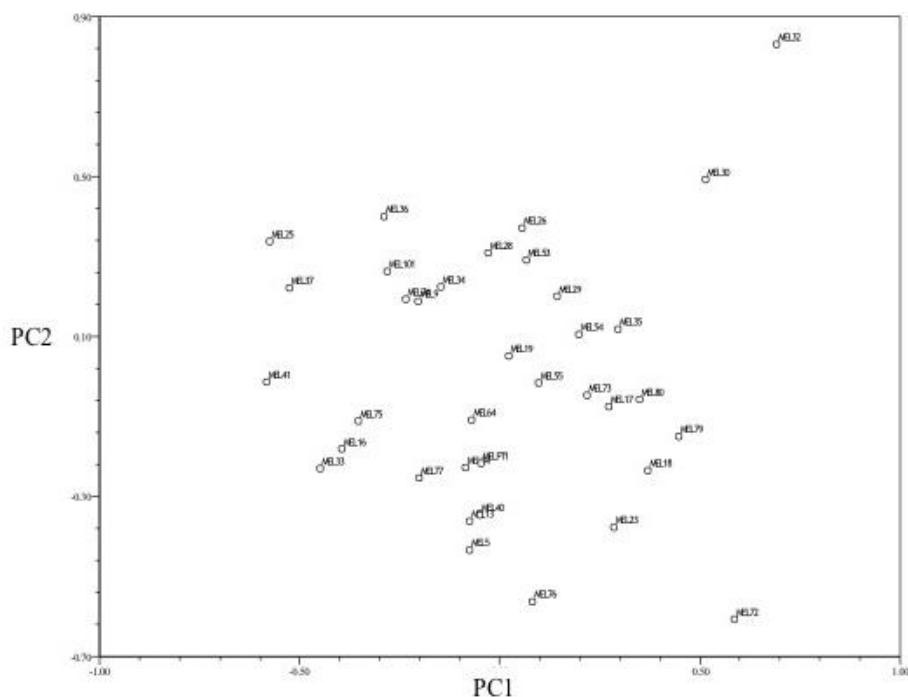
  

Sequenciação 3	Familia	Espécie	% Sequências
		<i>Trigonella cretica</i>	37,42
Fabaceae		<i>Ulex sp.</i>	22,35
		<i>Medicago sp.</i>	11,26
	Rosaceae	<i>Prunus sp.</i>	16,70
Rhamnaceae		<i>Rhamnus sp.</i>	5,61
Oxalidaceae		<i>Oxalis sp.</i>	3,80
Myrtaceae		<i>Eucalyptus sp.</i>	2,86

Estes resultados enfatizam como esta técnica permite a obtenção de resultados de confiança e qualidade uma vez que estes são precisos e reprodutíveis.

#### 4.4. Origem Geográfica

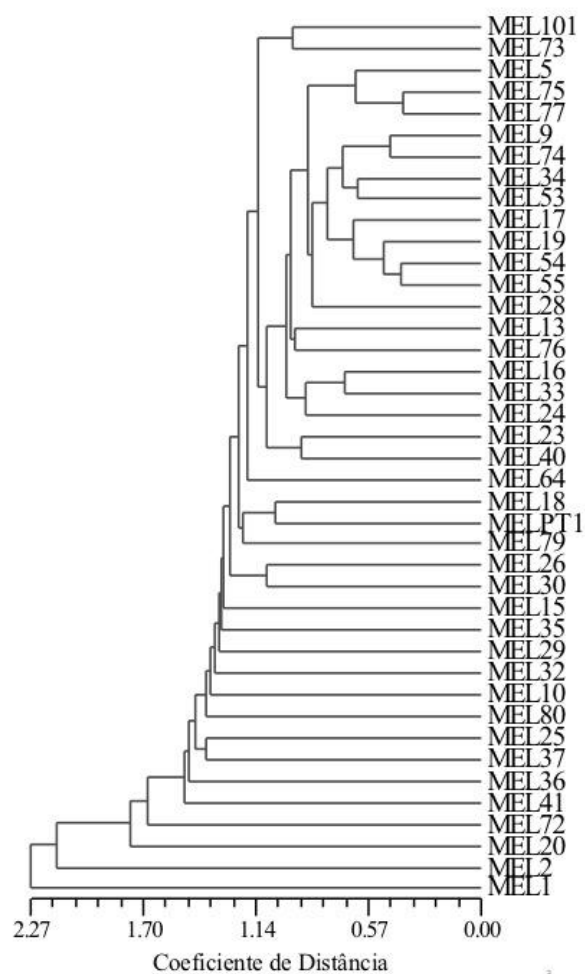
De maneira a averiguar se existia uma relação entre as famílias identificadas através da sequenciação e o local de produção de cada amostra de mel, procedeu-se à construção de um dendrograma com base na distância euclidiana. Para tal, construiu-se uma matriz com as famílias identificadas em todas as amostras, realizou-se uma análise de componentes principais, calculou-se a distância euclidiana entre cada uma e elaborou-se o dendrograma. A análise de componentes principais permitiu avaliar a variância entre as amostras de mel (Figura 4.6). O gráfico de componentes principais mostrou que maioria das amostras se encontra agrupada num só grupo. Isto significa que a variância entre amostras é tão elevada, ou seja, que as famílias identificadas são variáveis entre amostras, que não é possível agrupá-las em grupos distintos.



**Figura 4.6** Gráfico de componentes principais com as diferentes amostras de mel.

Construiu-se um dendrograma de distância com base na distância euclidiana (Figura 4.7). Mais uma vez, foi possível observar que as amostras de mel estudadas apresentam uma grande variância entre eles. Desde logo a primeira separação revela que a amostra de Mel 1 é muito diferente de todas as outras, seguido da separação do Mel 2 das restantes, ocorrendo sucessivamente separações individuais das amostras 20, 72, 41 e 36 de todas as outras amostras.



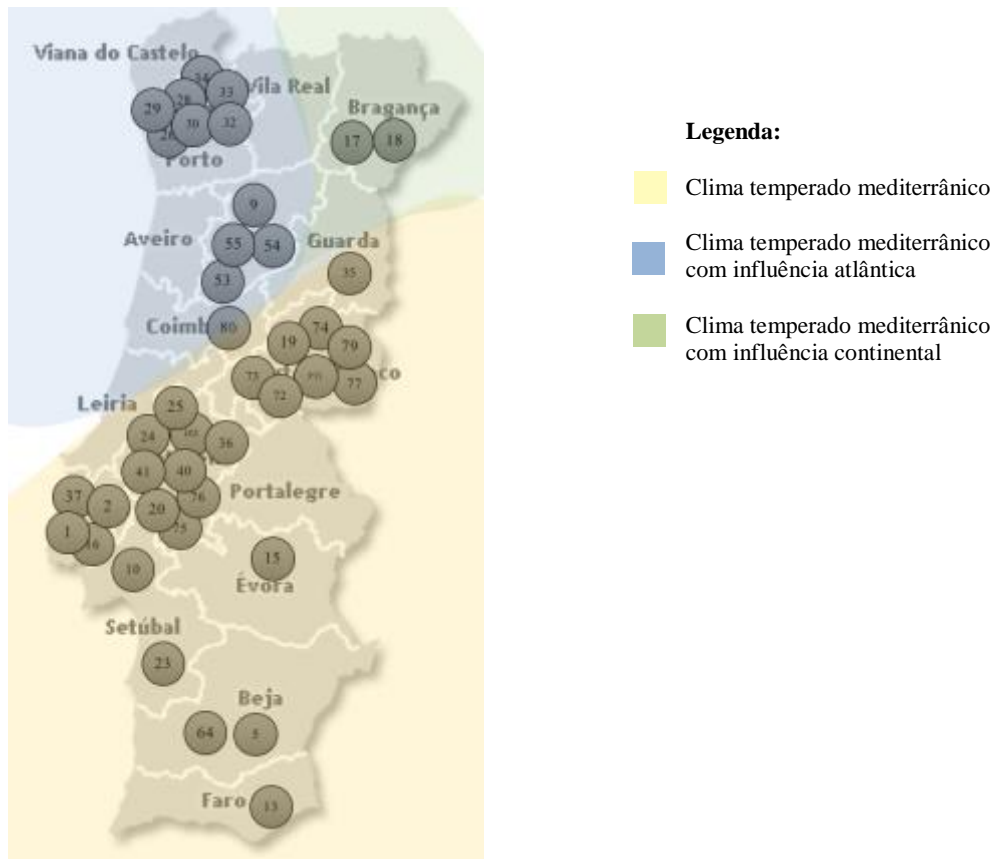


**Figura 4.7** Dendrograma baseado na distância euclidiana.

Embora não seja possível estabelecer uma relação direta entre o local de produção das amostras de mel e as famílias botânicas identificadas por NGS, esta técnica permitiu identificar plantas características da flora local.

A flora é diretamente influenciada pelo clima, e com isto, constatou-se que as amostras de mel estudadas apresentam grãos de pólen de plantas endêmicas de essencialmente duas regiões geográficas de Portugal (Figura 4.8):

- i. região de clima temperado mediterrânico;
- ii. região de clima temperado mediterrânico com influência atlântica.

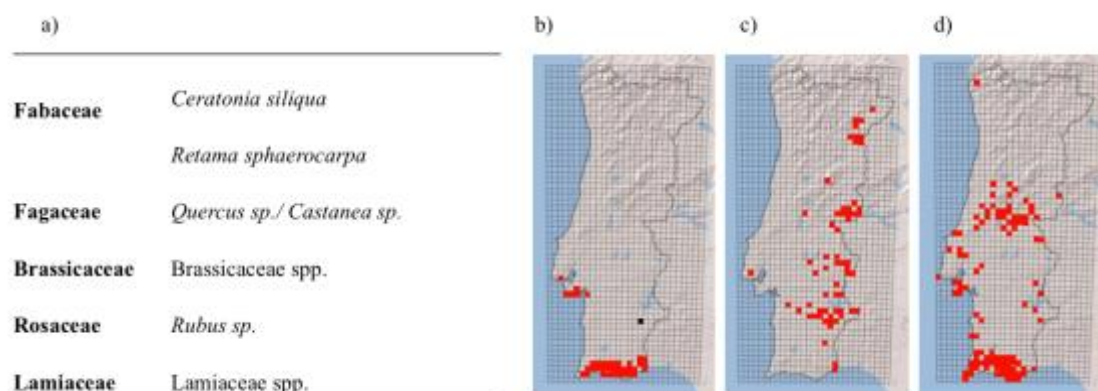


**Figura 4.8** Origem geográfica das amostras de mel.

A maioria das amostras estudadas provêm de uma zona de clima temperado mediterrânico.

Esta região é caracterizada pela baixa pluviosidade, o que permite a existência de arbustos e árvores de folha pequena e persistente como o sobreiro (*Quercus suber*), azinheira (*Quercus rotundifolia*), tojo-do-sul (*Genista hirsuta*), rosmaninho (*Lavandula stoechas*), alfarroba (*Ceratonia siliqua*), entre outros <sup>49,50</sup>.

Em alguns casos, as espécies botânicas identificadas por NGS das amostras localizadas em zonas de clima temperado mediterrânico mostraram ser coincidentes com algumas espécies pertencentes à flora local – como é o caso da alfarroba (*Ceratonia siliqua*) e do piorno (*Retama sphaerocarpa*), ambas espécies endémicas (Figura 4.9).



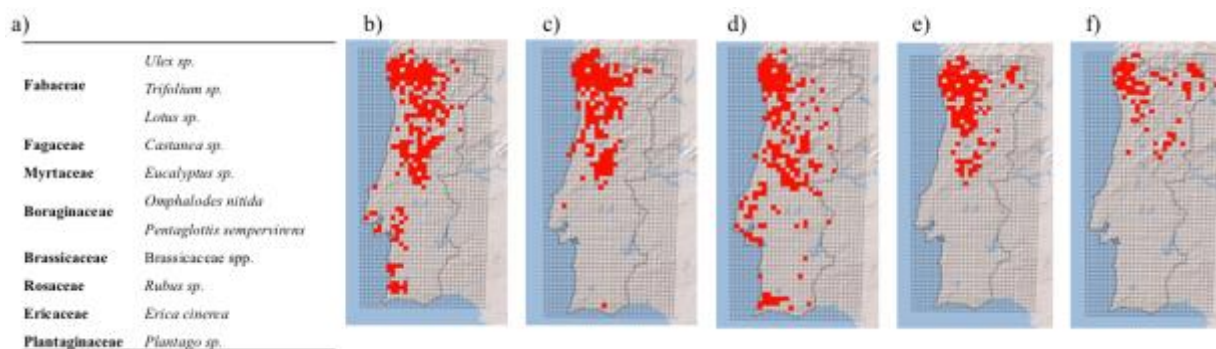
**Figura 4.9** Espécies florais identificadas por NGS de uma amostra de mel oriunda de uma zona de clima temperado mediterrânico. a) Resultado da sequenciação; b) – d) Mapa de distribuição de *Ceratonia siliqua*<sup>51</sup>, *Retama sphaerocarpa*<sup>52</sup> e *Lavandula stoechas*<sup>53</sup> respetivamente. Adaptado de: Flora-on (<http://flora-on.pt/>).

A análise melissopalinológica de algumas amostras provenientes desta zona geográfica identificou grãos de pólen de rosmaninho (*Lavandula stoechas*) – espécie esta que é característica da zona de clima temperado mediterrânico.

Como foi mencionado anteriormente, as sequências das regiões-alvo do rosmaninho aqui usadas não estão disponíveis na base de dados, portanto a sua identificação não é possível. No entanto, a sequenciação por NGS possibilitou a identificação de DNA pertencente à família das lamiáceas. Para comprovar se as sequências pertencentes à família Lamiaceae corresponderiam a *Lavandula stoechas*, procedeu-se ao alinhamento das sequências obtidas com a sequência de *L. stoechas* conseguida através da sequenciação do DNA extraído diretamente do rosmaninho. Verificou-se as sequências apresentavam um grau de identidade elevado.

Das 41 amostras, 14 localizam-se na zona de clima temperado mediterrânico com influência atlântica.

Esta zona do país tem um clima temperado e chuvoso. O clima deste território permite a presença de plantas da denominada "flora atlântica" como os carvalhos (*Quercus sp.*), tojais (*Ulex sp.*) e urzes (*Erica sp.*)<sup>49</sup>. Nas últimas décadas e devido à ação humana, esta zona tem vindo a ser reflorestada com eucaliptos (*Eucalyptus sp.*). Estes, agora, também caracterizam a flora desta zona de Portugal<sup>54</sup>. Os resultados obtidos mostram que a identificação das famílias e espécies botânicas das amostras de mel provenientes da zona com influência atlântica, é concordante com a flora típica dessa mesma zona (Figura 4.10).



**Figura 4.10** Espécies florais identificadas por NGS de uma amostra de mel proveniente de uma zona de clima temperado mediterrânico com influência atlântica. a) Resultado da sequenciação; b) – f) Mapa de distribuição de *Ulex minor*<sup>55</sup>, *Erica cinerea*<sup>56</sup>, *Eucalyptus* sp.<sup>57</sup>, *Omphalodes nitida*<sup>58</sup> e *Pentaglottis sempervirens*<sup>59</sup>, respetivamente. Adaptado de: Flora-on (<http://flora-on.pt/>).

É de salientar que, apesar de não se poder estabelecer uma relação direta entre os locais de produção das amostras de mel e a identificação das famílias por NGS, esta técnica pode ser aplicada de maneira a identificar plantas características da flora do local de proveniência do mel. Isto poderá contribuir para a garantia de autenticidade do produto e também em casos em que é necessário rastrear a sua origem.

Posto isto, conclui-se que o melhoramento da técnica de NGS conjuntamente com a superação das limitações associadas, permitirá a sua utilização como técnica fiável para determinar a origem geográfica e botânica de amostras de mel, assegurando quer a qualidade e autenticidade do produto em questão, quer a confiança do consumidor.

## 5. Conclusão e Perspetivas Futuras

O presente estudo teve como objetivo caracterizar 41 amostras de mel através de técnicas de NGS. Para tal, definiu-se um pré-tratamento inicial das amostras e um método de extração de DNA dos grãos de pólen presentes no mel. O DNA extraído foi amplificado e sequenciado com dois pares de *primers* específicos que possibilitam identificação da origem botânica.

Conclui-se que, em geral, esta técnica de sequenciação permite a identificação de famílias e/ou espécies botânicas presentes no mel. Embora não seja possível, para já, fazer uma classificação quantitativa das espécies botânicas, o facto de se poder neste momento identificar as famílias que o constituem é, em si, um avanço notável e um bom ponto de partida para uma investigação futura. No caso de fraude alimentar, é imperativo identificar as famílias botânicas que constituem o mel e essa identificação foi possível para a maioria dos casos neste trabalho.

Num futuro próximo, a técnica NGS poderá vir a ser aplicada de modo a garantir de forma rigorosa e fiável a autenticidade do produto através da determinação da sua origem botânica e geográfica. Para que esse futuro possa ser uma realidade a curto prazo, é necessário ultrapassar primeiro as limitações existentes, tais como a atualização das bases de dados com sequências das regiões propostas para identificação de espécies, ou mesmo a criação de novas bases de dados. A superação desta lacuna poderá permitir que a técnica NGS se torne o novo *gold-standard* para a caracterização de amostras de mel. Esta técnica, apesar de ainda necessitar alguma otimização, é extremamente vantajosa em relação à melissopalinologia, uma vez que permite obter uma identificação num curto espaço de tempo.

Adicionalmente, a determinação da origem botânica e geográfica do mel permite conhecer a visitação floral das abelhas, assegurando a qualidade do mel. Para além da identificação da origem botânica, é ainda possível alargar a técnica de NGS à identificação de microrganismos, como fungos e bactérias presentes no mel.

## 6. Bibliografia

1. Codex Stan 12-1981. Codex Standard for Honey. 1–8 (2001).
2. Union, E. Council Directive 2001/110/EC relating to honey. *Off. J. Eur. Communities* 10–47 (2002).
3. Wheeler, M. M. & Robinson, G. E. Diet-dependent gene expression in honey bees: honey vs. sucrose or high fructose corn syrup. *Sci. Rep.* **4**, 5726 (2014).
4. White, J. W. & Doner, L. W. Honey Composition and Properties. *Agric. Handb.* **335**, 82–91 (1980).
5. Da Silva, P. M., Gauche, C., Gonzaga, L. V., Costa, A. C. O. & Fett, R. Honey: Chemical composition, stability and authenticity. *Food Chem.* **196**, 309–323 (2016).
6. Bogdanov, S., Tomislav, J., Sieber, R. & Gallmann, P. Honey for Nutrition and Health: a Review. *Am. J. Coll. Nutr.* **27**, 677–689 (2008).
7. Kowalski, S., Łukasiewicz, M. & Berski, W. Applicability of physico-chemical parameters of honey for identification of the botanical origin. *Acta Sci. Pol. Technol. Aliment.* **12**, 51–59 (2013).
8. Louveaux, J., Maurizio, A. & Vorwohl, G. Methods of melissopalynology. *Bee World* **5**, 139–153 (1978).
9. Schievano, E., Finotello, C., Uddin, J., Mammi, S. & Piana, L. Objective Definition of Monofloral and Polyfloral Honeys Based on NMR Metabolomic Profiling. *J. Agric. Food Chem.* acs.jafc.6b00619 (2016).
10. Von Der Ohe, W., Persano Oddo, L., Piana, M. L., Morlot, M. & Martin, P. Harmonized methods of melissopalynology. *Apidologie* **35**, S18–S25 (2004).
11. Bogdanov, S. Biological and therapeutic properties of bee venom. 1–21 (2014).
12. Mărghitaș, L. A. *et al.* In vitro antioxidant capacity of honeybee-collected pollen of selected floral origin harvested from Romania. *Food Chem.* **115**, 878–883 (2009).
13. Viuda-Martos, M., Ruiz-Navajas, Y., Fernández-López, J. & Pérez-Álvarez, J. A. Functional Properties of Honey, Propolis, and Royal Jelly. *J. Food Sci.* **73**, R117–R124 (2008).
14. McLoone, P., Warnock, M. & Fyfe, L. Honey: A realistic antimicrobial for disorders of the skin. *J. Microbiol. Immunol. Infect.* **49**, 161–167 (2016).
15. Erejuwa, O. O., Sulaiman, S. A. & Ab Wahab, M. S. Honey - A novel antidiabetic agent. *Int. J. Biol. Sci.* **8**, 913–934 (2012).
16. Soares, S., Amaral, J. S., Oliveira, M. B. P. P. & Mafra, I. Improving DNA isolation from honey for the botanical origin identification. *Food Control* **48**, 130–136 (2015).
17. Moore, J. C., Spink, J. & Lipp, M. Development and Application of a Database of Food Ingredient Fraud and Economically Motivated Adulteration from 1980 to 2010. *J. Food Sci.* **77**, (2012).
18. Committee on the Environment Public Health and Food Safety. Report on the food crisis, fraud in the food chain and the control thereof (2013/2091(INI)).

- (2013).
19. Spink, J. & Moyer, D. C. Defining the Public Health Threat of Food Fraud. *J. Food Sci.* **76**, (2011).
  20. Manning, L. & Soon, J. M. Food Safety, Food Fraud, and Food Defense: A Fast Evolving Literature. *J. Food Sci.* **81**, R823–R834 (2016).
  21. Gossner, C. M. E. *et al.* The melamine incident: Implications for international food and feed safety. *Environ. Health Perspect.* **117**, 1803–1808 (2009).
  22. Dorne, J. Lou *et al.* Recent advances in the risk assessment of melamine and cyanuric acid in animal feed. *Toxicol. Appl. Pharmacol.* **270**, 218–229 (2013).
  23. O'Mahony, P. J. Finding horse meat in beef products--a global problem. *QJM* **106**, 595–597 (2013).
  24. Hawkins, J. *et al.* Using DNA Metabarcoding to Identify the Floral Composition of Honey: A New Tool for Investigating Honey Bee Foraging Preferences. *PLoS One* **10**, e0134735 (2015).
  25. Bruni, I. *et al.* A DNA barcoding approach to identify plant species in multiflower honey. *Food Chem.* **170**, 308–315 (2015).
  26. Valentini, A., Miquel, C. & Taberlet, P. DNA barcoding for honey biodiversity. *Diversity* **2**, 610–617 (2010).
  27. Keller, A. *et al.* Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biol.* **in press**, 558–566 (2014).
  28. Kraaijeveld, K. *et al.* Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Mol. Ecol. Resour.* (2014).
  29. Galimberti, A. *et al.* A DNA Barcoding Approach to Characterize Pollen Collected by Honeybees. *PLoS One* **9**, e109363 (2014).
  30. Zhou, X. *et al.* The next-generation sequencing technology and application. *Protein Cell* **1**, 520–36 (2010).
  31. Kohn, A. B., Moroz, T. P., Barnes, J. P., Netherton, M. & Moroz, L. L. in **1048**, 247–284 (2013).
  32. Mardis, E. R. Next-Generation Sequencing Platforms. *Annu. Rev. Anal. Chem.* **6**, 287–303 (2013).
  33. Van Dijk, E. L., Auger, H., Jaszczyszyn, Y. & Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet.* **30**, (2014).
  34. Lalmangaihi, R., Ghatak, S., Laha, R., Gurusubramanian, G. & Kumar, N. S. Protocol for Optimal Quality and Quantity Pollen DNA Isolation from Honey Samples. *J. Biomol. Tech.* jbt.14-2504-001 (2014).
  35. Bell, K. L. *et al.* Pollen DNA barcoding : current applications and future. **640**, 629–640 (2016).
  36. Jain, S. & Jesus, F. Extraction of DNA from honey and its amplification by PCR for botanical identification. *Food Sci. Technol.* **2013**, 753–756 (2013).
  37. Guertler, P., Eicheldinger, A., Muschler, P., Goerlich, O. & Busch, U. Automated DNA extraction from pollen in honey. *Food Chem.* **149**, 302–306 (2014).
  38. Barbosa, C., Nogueira, S., Gadanho, M. & Chaves, S. in *Molecular Microbial*

- Diagnostic Methods: Pathways to Implementation in the Food and Water Industry* 135–154 (Elsevier Inc., 2016). doi:10.1016/B978-0-12-416999-9.00007-1
39. Rohlf, F. J. *NTSYSpc: Numerical Taxonomy and Multivariate Analysis System. Ver 2.21.* (Exeter Software, 2009).
  40. Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* **17**, 1105–1109 (1991).
  41. *Genomic DNA from food - User manual NucleoSpin Food.* (Macherey-Nagel, 2014).
  42. *E. Z. N. A. ® Tissue DNA Kit.* (Omega Bio-tek, 2013).
  43. *DNeasy® mericon™ Food Handbook - for extraction of total nucleic acids from a range of food sample types.* (2010).
  44. *E. Z. N. A. ® SP Plant DNA Midi Kit.* (Omega Bio-tek, 2013).
  45. *DNeasy Plant Handbook.* (Qiagen, 2012).
  46. Aizenberg-Gershtein, Y., Izhaki, I. & Halpern, M. Do Honeybees Shape the Bacterial Community Composition in Floral Nectar? *PLoS One* **8**, (2013).
  47. Sinacori, M. *et al.* Cultivable microorganisms associated with honeys of different geographical and botanical origin. *Food Microbiol.* **38**, 284–294 (2014).
  48. Olaitan, P. B., Adeleke, O. E. & Ola, I. O. Honey: A reservoir for microorganisms and an inhibitory agent for microbes. *Afr. Health Sci.* **7**, 159–165 (2007).
  49. Aguiar, C., Mesquita, S. & Honrado, J. Biogeografia e uso do território. *Atlas das Aves Nidificantes em Port.* 41–47 (1999).
  50. Costa, J. C., Aguiar, C., Capelo, J. H., Lousã, M. & Neto, C. Biogeografia de Portugal Continental. *Quercetea* **0**, 5–56 (1998).
  51. Carapeto, A. *et al.* *Ceratonia siliqua* L. - mapa de distribuição. *Flora-On Flora Port. Interactiva, Soc. Port. Botânica.* <http://www.flora-on.pt/#wCeratonia+siliqua>. Consult. Realiz. em 5/08/2016
  52. Porto, M. *et al.* *Retama sphaerocarpa* (L.) Boiss. - mapa de distribuição. *Flora-On Flora Port. Interactiva, Soc. Port. Botânica.* <http://www.flora-on.pt/#wRetama+sphaerocarpa>. Consult. Realiz. em 5/08/2016
  53. Carapeto, A. *et al.* *Lavandula stoechas* L. subsp. *stoechas* - mapa de distribuição. *Flora-On Flora Port. Interactiva, Soc. Port. Botânica.* <http://www.flora-on.pt/#wLavandula+stoechas+subsp.+stoechas>. Consult. Realiz. em 5/08/2016
  54. Alves, A. M., Pereira, J. S. & Silva, J. M. N. in *Pinhais e eucaliptais - A floresta cultivada. Coleção Árvores e Florestas de Portugal* **4**, 151–165 (2007).
  55. Lourenço, J. *et al.* *Ulex minor* Roth - mapa de distribuição. *Flora-On Flora Port. Interactiva, Soc. Port. Botânica.* <http://www.flora-on.pt/#wUlex+minor>. Consult. Realiz. em 4/08/2016
  56. Araújo, P. V. *et al.* *Erica cinerea* L. - mapa de distribuição. *Flora-On Flora Port. Interactiva, Soc. Port. Botânica.* <http://www.flora-on.pt/#wErica+cinerea>. Consult. Realiz. em 4/08/2016



57. Lourenço, J. *et al.* *Eucalyptus globulus* Labill. - mapa de distribuição. *Flora-On Flora Port. Interactiva, Soc. Port. Botânica*. <http://www.flora-on.pt/#wEucalyptus+globulus>. Consult. Realiz. em 4/08/2016
58. Araújo, P. V. *et al.* *Omphalodes nitida* Hoffmanns. & Link - mapa de distribuição. *Flora-On Flora Port. Interactiva, Soc. Port. Botânica*. <http://www.flora-on.pt/#wOmphalodes+nitida>. Consult. Realiz. em 4/08/2016
59. Lourenço, J. *et al.* *Pentaglottis sempervirens* (L.) L.H.Bailey - mapa de distribuição. *Flora-On Flora Port. Interactiva, Soc. Port. Botânica*. <http://www.flora-on.pt/#wPentaglottis+sempervirens>. Consult. Realiz. em 4/08/2016

A.Anexos

Tabela A.1 Número médio de sequências obtidas e de famílias de plantas identificadas em 41 amostras de mel pela técnica de NGS.

Média do N° de Famílias Identificadas					Média do N° de Sequências				
Alvo					Alvo				
ribossomal					ribossomal				
plastidial					plastidial				
Amostra	ribossomal	plastidial	ribossomal	plastidial	Amostra	ribossomal	plastidial	ribossomal	plastidial
MEL 1	5	11	10555	9307	MEL 34	8	8	5675	5885
MEL 2	3	12	786	5900	MEL 35	7	7	3108	18147
MEL 5	2	5	7546	5661	MEL 36	5	-	4175	-
MEL 9	5	-	2181	-	MEL 37	1	6	17732	10294
MEL 10	3	10	6719	1860	MEL 40	5	9	4843	4289
MEL 13	3	8	5710	1665	MEL 41	8	4	1918	3516
MEL 15	5	10	12891	5102	MEL 53	10	5	6401	5373
MEL 16	3	-	1797	-	MEL 54	6	8	2265	7351
MEL 17	6	7	11716	5728	MEL 55	3	8	7575	5737
MEL 18	7	8	6469	4354	MEL 64	4	6	4052	11277
MEL 19	5	9	5897	3562	MEL 72	7	9	2457	14394
MEL 20	1	12	8396	2218	MEL 73	5	5	2666	12576
MEL 23	5	10	16527	5379	MEL 74	5	-	2649	-
MEL 24	2	4	5155	399	MEL 75	1	-	14713	-
MEL 25	6	5	9559	14384	MEL 76	1	5	865	4992
MEL 26	6	5	7498	15829	MEL 77	1	3	10200	22554
MEL 28	8	6	6542	7346	MEL 79	4	9	6062	6300
MEL 29	6	8	7551	5028	MEL 80	5	7	6385	2567
MEL 30	9	10	3138	5972	MEL PT1	3	6	5334	8001
MEL 32	13	9	93	8944	MEL 101	5	5	2573	10009
MEL 33	5	-	7310	-					

Score	Expect	Identities	Gaps	Strand
252 bits(136)	2e-63	139/140(99%)	1/140(0%)	Plus/Plus
Query 1	AAGCTGTTCTAACAAATGGAGTTGGTAGAGGAATCCTTCTATCGAAACTTACG	60		
Sbjct 102	AAGCTGTTCTAACAAATGGAGTTGGTAGAGGAATCCTTCTATCGAAACTTACG-AAAAAA	160		
Query 61	TGAAAGATAACTCTATATACATATAGTATTAGGTACTGAAATACTATATCTTATGATTAA	120		
Sbjct 161	TGAAAGATAACTCTATATACATATAGTATTAGGTACTGAAATACTATATCTTATGATTAA	220		
Query 121	TAAGAACAAAAATAGGTATT	140		
Sbjct 221	TAAGAACAAAAATAGGTATT	240		

Figura A.1 Resultado do alinhamento de uma sequência nucleotídica de uma região plastidial de Erica umbellata com a ferramenta BLAST.

Melissopalínologia		NGS	
Família	Espécie	Família	Espécie
Myrtaceae	<i>Eucalyptus sp.</i>	Myrtaceae	<i>Eucalyptus sp.</i>
Rhamnaceae	<i>Rhamnus sp.</i>	Rhamnaceae	<i>Rhamnus sp.</i>
Fabaceae	Fabaceae spp.	Fabaceae	Fabaceae spp.
Salicaceae	<i>Salix sp.</i>	Salicaceae	<i>Salix sp.</i>
Brassicaceae	<i>Raphanus raphanistrum</i>	Brassicaceae	Brassicaceae spp.
		Ericaceae	<i>Erica sp.</i>
		Lamiaceae	<i>Rosmarinus officinalis</i>

**Figura A.2** Imagem ilustrativa de famílias de plantas de uma amostra de mel que não foram identificadas pela melissopalínologia (evidenciadas por uma seta).

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

[Alignments](#) [Download](#) [GenBank](#) [Graphics](#) [Distance tree of results](#)

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	<a href="#">Eucalyptus erythrocorys chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180799.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus guilfoylei chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180798.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus salmoneophloia chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180796.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus diversicolor chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180795.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus torquata chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180794.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus spathulata chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180793.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus deglupta chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180792.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus camaldulensis chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180791.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus saligna chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180790.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus aromaphloia chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180789.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus nitens chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180788.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus globulus chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180787.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus diadocalyx chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180786.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus polybractea chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180785.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus melliodora chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180783.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus marginata chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180781.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus cloeziana chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180779.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus umbra chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180778.1</a>
<input type="checkbox"/>	<a href="#">Eucalyptus regnans chloroplast, complete genome</a>	255	255	98%	1e-64	100%	<a href="#">KC180777.1</a>

**Figura A.3** Imagem representativa do resultado do alinhamento das sequências nucleotídicas com a ferramenta BLAST e da falta de poder de resolução dos *primers* utilizados na identificação da espécie *Eucalyptus sp.*

**Tabela A.2.** Famílias dos grãos de pólen predominantes e secundários identificados através da melissopalínologia e NGS. Quando a análise melissopalínológica das amostras de mel identificou as mesmas famílias de plantas que a análise das sequências, usou-se um ✓. Caso contrário, usou-se um ✗.

Amostra	Família	✓/✗	Amostra	Família	✓/✗	Amostra	Família	✓/✗	Amostra	Família	✓/✗
101	Ericaceae	✓	MEL 16	Fabaceae	✓	MEL 26	Myrtaceae	✓	MEL 36	Rosaceae	✓
	Lamiaceae	✓		Myrtaceae	✓		Fagaceae	✓		Myrtaceae	✓
	Rosaceae	✓		Rosaceae	✓		Fabaceae	✓		Brassicaceae	✓
MEL 1	Myrtaceae	✓	MEL 17	Fagaceae	✓	MEL 28	Myrtaceae	✓	MEL 37	Myrtaceae	✓
	Malvaceae	✓		Rosaceae	✓		Brassicaceae	✓		Rhamnaceae	✓
	Rosaceae	✓		Lamiaceae	✓		Rosaceae	✓		Fabaceae	✓
MEL 2	Myrtaceae	✓	MEL 18	Fagaceae	✓	MEL 29	Myrtaceae	✓	MEL 40	Ericaceae	✓
	Fabaceae	✓		Rosaceae	✓		Fagaceae	✓		Lamiaceae	✓
	Asteraceae	✓		Lamiaceae	✓		Rosaceae	✓		Myrtaceae	✓
MEL 5	Lamiaceae	✓	MEL 19	Ericaceae	✓	MEL 30	Myrtaceae	✓	MEL 41	Myrtaceae	✓
	Fagaceae	✓		Fabaceae	✓		Fagaceae	✓		Rhamnaceae	✓
	Fabaceae	✓		Rosaceae	✓		Boraginaceae	✓		Fabaceae	✓
MEL 9	Myrtaceae	✓	MEL 20	Fabaceae	✓	MEL 32	Myrtaceae	✓	MEL 53	Myrtaceae	✓
	Fabaceae	✓		Lamiaceae	✓		Fagaceae	✓		Fagaceae	✓
	Rosaceae	✓		Fagaceae	✓		Boraginaceae	✓		Fabaceae	✓
MEL 10	Ericaceae	✓	MEL 23	Boraginaceae	✓	MEL 33	Ericaceae	✓	MEL 54	Fagaceae	✓
	Lamiaceae	✓		Rosaceae	✓		Fabaceae	✓		Myrtaceae	✓
	Fabaceae	✓		Fabaceae	✓		Asteraceae	✓		Rosaceae	✓
MEL 13	Lamiaceae	✓	MEL 24	Fabaceae	✓	MEL 34	Myrtaceae	✓	MEL 55	Rosaceae	✓
	Ericaceae	✓		Lamiaceae	✓		Fagaceae	✓		Myrtaceae	✓
	Fabaceae	✓		Asteraceae	✓		Rosaceae	✓		Fagaceae	✓
MEL 15	Fabaceae	✓	MEL 25	Myrtaceae	✓	MEL 35	Fagaceae	✓	MEL 64	Myrtaceae	✓
	Asteraceae	✓		Fabaceae	✓		Rosaceae	✓		Rosaceae	✓
	Rosaceae	✓		Rhamnaceae	✓		Fabaceae	✓		Rosaceae	✓
										MEL PT1	✓
										Lamiaceae	✓

